WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Jan 2, 2024 – 11:55 pm GMT

PDB ID	:	5FLV
Title	:	Crystal structure of NKX2-5 and TBX5 bound to the Nppa promoter region
Authors	:	Stirnimann, C.U.; Glatt, S.; Mueller, C.W.
Deposited on	:	2015-10-28
Resolution	:	3.00 Å(reported)

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

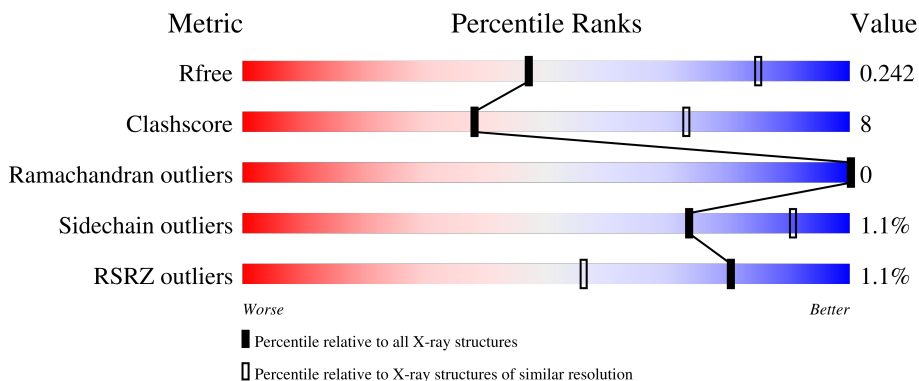|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 3.00 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.

Metric	Percentile Ranks	Value
Rfree		0.242
Clashscore		8
Ramachandran outliers		0
Sidechain outliers		1.1%
RSRZ outliers		1.1%

Worse — *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	285	67% 13% 21%
1	E	285	62% 17% 20%
1	I	285	63% 16% 21%
1	M	285	64% 16% 20%
2	B	22	59% 36% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	22	77% 18% 5%
2	J	22	64% 32% 5%
2	N	22	59% 36% 5%
3	C	22	64% 27% 9%
3	G	22	50% 45% 5%
3	K	22	64% 36%
3	O	22	59% 41%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 11035 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 HOMEOBOX PROTEIN NKX-2.5, T-BOX TRANSCRIPTION FACTOR TBX5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	226	Total	C	N	O	S	0	0	0
			1872	1211	339	316	6			
1	E	227	Total	C	N	O	S	0	0	0
			1893	1223	346	318	6			
1	I	226	Total	C	N	O	S	0	0	0
			1884	1215	345	318	6			
1	M	229	Total	C	N	O	S	0	0	0
			1901	1230	344	321	6			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	130	GLY	-	expression tag	UNP P42582
A	131	ALA	-	expression tag	UNP P42582
A	132	MET	-	expression tag	UNP P42582
A	133	GLY	-	expression tag	UNP P42582
A	192	SER	CYS	engineered mutation	UNP P42582
A	198	SER	-	linker	UNP P42582
A	199	SER	-	linker	UNP P42582
A	200	SER	-	linker	UNP P42582
A	201	SER	-	linker	UNP P42582
A	202	SER	-	linker	UNP P42582
A	203	SER	-	linker	UNP P42582
A	204	SER	-	linker	UNP P42582
A	205	SER	-	linker	UNP P42582
A	206	SER	-	linker	UNP P42582
A	207	SER	-	linker	UNP P42582
A	208	SER	-	linker	UNP P42582
A	209	SER	-	linker	UNP P42582
A	210	SER	-	linker	UNP P42582
A	211	SER	-	linker	UNP P42582
A	212	SER	-	linker	UNP P42582

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	213	ALA	-	linker	UNP P42582
A	1202	SER	CYS	engineered mutation	UNP P70326
E	130	GLY	-	expression tag	UNP P42582
E	131	ALA	-	expression tag	UNP P42582
E	132	MET	-	expression tag	UNP P42582
E	133	GLY	-	expression tag	UNP P42582
E	192	SER	CYS	engineered mutation	UNP P42582
E	198	SER	-	linker	UNP P42582
E	199	SER	-	linker	UNP P42582
E	200	SER	-	linker	UNP P42582
E	201	SER	-	linker	UNP P42582
E	202	SER	-	linker	UNP P42582
E	203	SER	-	linker	UNP P42582
E	204	SER	-	linker	UNP P42582
E	205	SER	-	linker	UNP P42582
E	206	SER	-	linker	UNP P42582
E	207	SER	-	linker	UNP P42582
E	208	SER	-	linker	UNP P42582
E	209	SER	-	linker	UNP P42582
E	210	SER	-	linker	UNP P42582
E	211	SER	-	linker	UNP P42582
E	212	SER	-	linker	UNP P42582
E	213	ALA	-	linker	UNP P42582
E	1202	SER	CYS	engineered mutation	UNP P70326
I	130	GLY	-	expression tag	UNP P42582
I	131	ALA	-	expression tag	UNP P42582
I	132	MET	-	expression tag	UNP P42582
I	133	GLY	-	expression tag	UNP P42582
I	192	SER	CYS	engineered mutation	UNP P42582
I	198	SER	-	linker	UNP P42582
I	199	SER	-	linker	UNP P42582
I	200	SER	-	linker	UNP P42582
I	201	SER	-	linker	UNP P42582
I	202	SER	-	linker	UNP P42582
I	203	SER	-	linker	UNP P42582
I	204	SER	-	linker	UNP P42582
I	205	SER	-	linker	UNP P42582
I	206	SER	-	linker	UNP P42582
I	207	SER	-	linker	UNP P42582
I	208	SER	-	linker	UNP P42582
I	209	SER	-	linker	UNP P42582
I	210	SER	-	linker	UNP P42582

Continued from previous page...

| Chain	Residue	Modelled	Actual	Comment	Reference
I	211	SER	-	linker	UNP P42582
I	212	SER	-	linker	UNP P42582
I	213	ALA	-	linker	UNP P42582
I	1202	SER	CYS	engineered mutation	UNP P70326
M	130	GLY	-	expression tag	UNP P42582
M	131	ALA	-	expression tag	UNP P42582
M	132	MET	-	expression tag	UNP P42582
M	133	GLY	-	expression tag	UNP P42582
M	192	SER	CYS	engineered mutation	UNP P42582
M	198	SER	-	linker	UNP P42582
M	199	SER	-	linker	UNP P42582
M	200	SER	-	linker	UNP P42582
M	201	SER	-	linker	UNP P42582
M	202	SER	-	linker	UNP P42582
M	203	SER	-	linker	UNP P42582
M	204	SER	-	linker	UNP P42582
M	205	SER	-	linker	UNP P42582
M	206	SER	-	linker	UNP P42582
M	207	SER	-	linker	UNP P42582
M	208	SER	-	linker	UNP P42582
M	209	SER	-	linker	UNP P42582
M	210	SER	-	linker	UNP P42582
M	211	SER	-	linker	UNP P42582
M	212	SER	-	linker	UNP P42582
M	213	ALA	-	linker	UNP P42582
M	1202	SER	CYS	engineered mutation	UNP P70326

- Molecule 2 is a DNA chain called 5'-D(*TP*CP*TP*TP*CP*TP*CP*AP*CP*AP*CP*CP*TP*TP*TP*GP*AP*AP*GP*TP*GP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	21	Total 426	C 204	N 72	O 129	P 21	0	0	0
2	F	21	Total 426	C 204	N 72	O 129	P 21	0	0	0
2	J	21	Total 425	C 203	N 72	O 129	P 21	0	0	0
2	N	21	Total 422	C 203	N 72	O 127	P 20	0	0	0

- Molecule 3 is a DNA chain called 5'-D(*AP*CP*CP*AP*CP*TP*TP*CP*AP*AP*AP*GP*GP*TP*GP*TP*GP*AP*GP*AP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	20	Total	C	N	O	P	0	0	0
			413	196	80	117	20			
3	G	21	Total	C	N	O	P	0	0	0
			435	206	85	123	21			
3	K	22	Total	C	N	O	P	0	0	0
			456	216	90	128	22			
3	O	22	Total	C	N	O	P	0	0	0
			456	216	90	128	22			

- Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).
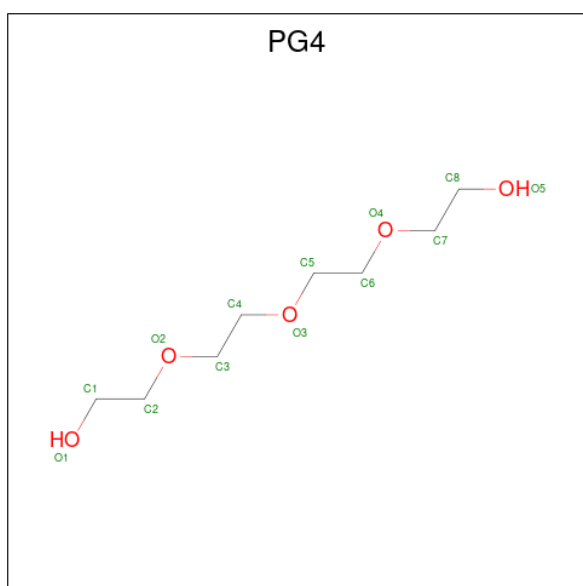
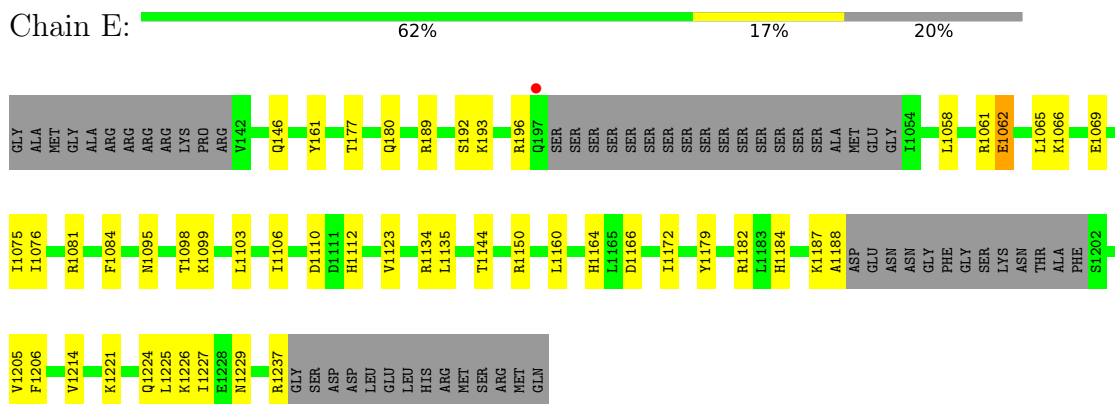| Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			13	8	5		
4	C	1	Total	C	O	0	0
			13	8	5		

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.
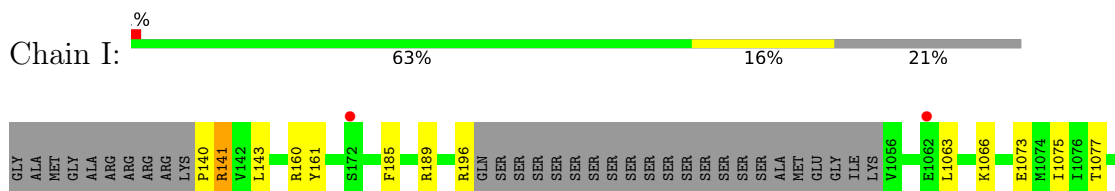
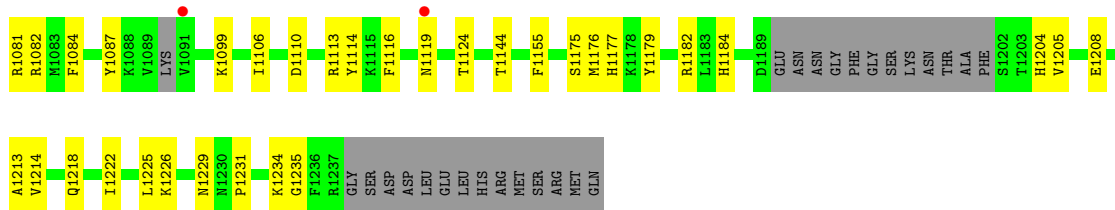• Molecule 1: HOMEOBOX PROTEIN NKX-2.5, T-BOX TRANSCRIPTION FACTOR TBX5

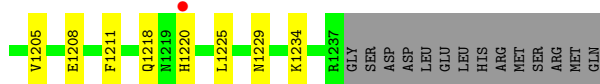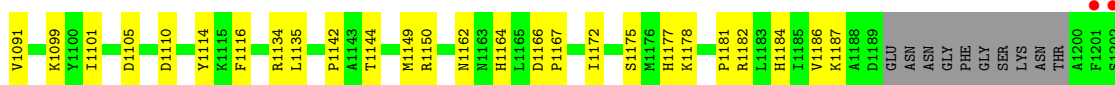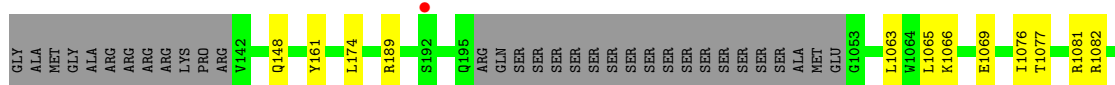• Molecule 1: HOMEOBOX PROTEIN NKX-2.5, T-BOX TRANSCRIPTION FACTOR TBX5

• Molecule 1: HOMEOBOX PROTEIN NKX-2.5, T-BOX TRANSCRIPTION FACTOR TBX5

R1081 R1082 M1083 F1084 Y1087 K1088 V1089 LYS V1091 K1099 I1106 D1110 R1113 Y1114 K1115 F1116 N1119 T1124 T1144 F1155 S1175 M1176 H1177 K1178 Y1179 R1182 L1183 H1184 D1189 GLU ASN ASN GLY PHE GLY SER LYS ASN THR ALA PHE S1202 T1203 H1204 V1205 E1208

A1213 V1214 Q1218 I1222 L1225 K1226 N1229 N1230 P1231 K1234 G1235 F1236 R1237 GLY SER ASP ASP LEU GLU LEU HIS ARG MET SER ARG MET GLN

• Molecule 1: HOMEOBOX PROTEIN NKX-2.5, T-BOX TRANSCRIPTION FACTOR TBX5

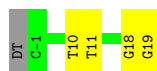Chain M:

| 64% | 16% | 20% |

GLY ALA MET GLY ALA ARG ARG ARG LYS PRO ARG V142 Q148 Y161 L174 R189 S192 Q195 ARG GLN SER SER SER SER SER SER SER SER SER SER SER ALA MET GLU G1053 L1063 W1064 L1065 K1066 E1069 I1076 T1077 R1081 R1082

V1091 K1099 Y1100 I1101 D1105 D1110 Y1114 K1115 F1116 R1134 L1135 P1142 A1143 T1144 M1149 R1150 N1162 N1163 H1164 D1166 P1167 I1172 S1175 M1176 H1177 K1178 P1181 R1182 L1183 H1184 I1185 V1186 K1187 A1188 D1189 GLU ASN ASN GLY PHE GLY SER LYS ASN THR A1200 F1201 S1202

V1205 E1208 F1211 Q1218 N1219 H1220 L1225 N1229 K1234 R1237 GLY SER ASP ASP LEU GLU LEU HIS ARG MET SER ARG MET GLN

• Molecule 2: 5'-D(*TP*CP*TP*TP*CP*TP*CP*AP*CP*AP*CP*CP*TP*TP *TP*GP*AP*A P*GP*TP*GP*G)-3'

Chain B:

| 59% | 36% | 5% |

DT C-1 C6 A7 T10 T11 T12 G13 A14 G18 G19

• Molecule 2: 5'-D(*TP*CP*TP*TP*CP*TP*CP*AP*CP*AP*CP*CP*TP*TP *TP*GP*AP*A P*GP*TP*GP*G)-3'

Chain F:

| 77% | 18% | 5% |

DT C-1 T10 T11 G18 G19

• Molecule 2: 5'-D(*TP*CP*TP*TP*CP*TP*CP*AP*CP*AP*CP*CP*TP*TP *TP*GP*AP*A P*GP*TP*GP*G)-3'

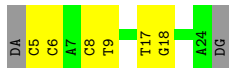Chain J:

| 64% | 32% | 5% |

DT C-1 C6 A7 C8 C9 T10 T11 G18 G19

• Molecule 2: 5'-D(*TP*CP*TP*TP*CP*TP*CP*AP*CP*AP*CP*CP*TP*TP *TP*GP*AP*A P*GP*TP*GP*G)-3'

Chain N:

• Molecule 3: 5'-D(*AP*CP*CP*AP*CP*TP*TP*CP*AP*AP*AP*GP*GP*TP *GP*TP*GP* AP*GP*AP*AP*G)-3'

Chain C:

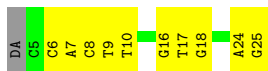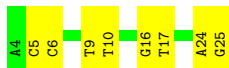• Molecule 3: 5'-D(*AP*CP*CP*AP*CP*TP*TP*CP*AP*AP*AP*GP*GP*TP *GP*TP*GP* AP*GP*AP*AP*G)-3'

Chain G:

• Molecule 3: 5'-D(*AP*CP*CP*AP*CP*TP*TP*CP*AP*AP*AP*GP*GP*TP *GP*TP*GP* AP*GP*AP*AP*G)-3'

Chain K:

• Molecule 3: 5'-D(*AP*CP*CP*AP*CP*TP*TP*CP*AP*AP*AP*GP*GP*TP *GP*TP*GP* AP*GP*AP*AP*G)-3'

Chain O:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	58.08Å 158.62Å 87.86Å 90.00° 89.95° 90.00°	Depositor
Resolution (Å)	19.97 – 3.00 19.97 – 3.01	Depositor EDS
% Data completeness (in resolution range)	98.0 (19.97-3.00) 97.9 (19.97-3.01)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.97 (at 2.98Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.190 , 0.241 0.192 , 0.242	Depositor DCC
R_{free} test set	1077 reflections (3.48%)	wwPDB-VP
Wilson B-factor (Å^2)	65.4	Xtriage
Anisotropy	0.705	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 37.5	EDS
L-test for twinning[2]	$< |L| > = 0.41, < L^2 > = 0.24$	Xtriage
Estimated twinning fraction	0.397 for h,-k,-l	Xtriage
Reported twinning fraction	0.570 for h,-k,-l	Depositor
Outliers	1 of 30939 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11035	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.79% 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 ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PG4

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	0.23	0/1923	0.41	0/2593
1	E	0.23	0/1943	0.41	0/2618
1	I	0.23	0/1934	0.40	0/2606
1	M	0.22	0/1952	0.39	0/2631
2	B	0.58	0/475	1.06	0/730
2	F	0.57	0/475	1.08	0/730
2	J	0.59	0/474	1.06	0/728
2	N	0.61	0/470	1.06	0/720
3	C	0.60	0/464	0.95	0/714
3	G	0.57	0/489	0.95	0/753
3	K	0.55	0/513	0.95	0/790
3	O	0.57	0/513	0.96	0/790
All	All	0.38	0/11625	0.69	0/16403

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts ⓘ

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	A	1872	0	1901	25	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1893	0	1930	35	0
1	I	1884	0	1909	28	0
1	M	1901	0	1930	29	0
2	B	426	0	239	4	0
2	F	426	0	239	3	0
2	J	425	0	236	6	0
2	N	422	0	238	7	0
3	C	413	0	225	4	0
3	G	435	0	236	6	0
3	K	456	0	247	7	0
3	O	456	0	247	11	0
4	A	13	0	18	0	0
4	C	13	0	18	0	0
All	All	11035	0	9613	146	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 8.

All (146) 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:J:19:DG:H1	3:K:5:DC:H5	1.37	0.72
1:I:1087:TYR:HB3	1:I:1155:PHE:HB2	1.71	0.72
1:E:1084:PHE:O	1:E:1226:LYS:NZ	2.22	0.69
1:E:177:THR:H	1:E:180:GLN:HE21	1.44	0.65
1:A:1099:LYS:HA	1:A:1144:THR:HA	1.77	0.65
1:A:167:ARG:HD2	1:A:178:SER:HB3	1.81	0.63
1:I:1084:PHE:O	1:I:1226:LYS:NZ	2.31	0.62
1:M:148:GLN:NE2	1:M:174:LEU:O	2.32	0.62
1:E:1182:ARG:HG2	1:E:1205:VAL:HG22	1.82	0.61
2:N:19:DG:N2	3:O:6:DC:C2	2.69	0.61
1:I:1099:LYS:HA	1:I:1144:THR:HA	1.83	0.60
1:M:1099:LYS:HA	1:M:1144:THR:HA	1.84	0.60
1:E:1187:LYS:HG2	1:E:1188:ALA:H	1.66	0.60
1:M:1082:ARG:HH11	3:O:16:DG:H5”	1.66	0.59
1:E:1182:ARG:HH21	1:E:1184:HIS:HE1	1.49	0.59
1:E:1225:LEU:O	1:E:1229:ASN:ND2	2.31	0.58
3:G:9:DT:H2”	3:G:10:DT:H5”	1.85	0.58
1:A:1103:LEU:HD22	1:E:1135:LEU:HD11	1.87	0.57
1:I:1231:PRO:HA	1:I:1234:LYS:HE3	1.86	0.57
1:E:1099:LYS:HA	1:E:1144:THR:HA	1.85	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:K:9:DT:H2"	3:K:10:DT:H5"	1.87	0.56
1:M:1234:LYS:HB3	2:N:9:DC:H4'	1.87	0.56
2:B:18:DG:H2"	2:B:19:DG:O5'	2.05	0.56
1:A:150:TYR:OH	1:E:1164:HIS:ND1	2.33	0.56
2:J:18:DG:H2"	2:J:19:DG:H8	1.72	0.55
1:I:1182:ARG:HH21	1:I:1184:HIS:CE1	2.25	0.55
1:E:1075:ILE:HD13	1:E:1214:VAL:HG23	1.88	0.55
1:E:1164:HIS:HA	1:E:1172:ILE:HG21	1.89	0.55
1:M:1164:HIS:HA	1:M:1172:ILE:HG21	1.90	0.54
1:E:161:TYR:HA	1:E:189:ARG:HH21	1.73	0.54
1:E:192:SER:O	1:E:196:ARG:HG3	2.08	0.54
2:J:6:DC:H2"	2:J:7:DA:C8	2.43	0.54
1:M:1175:SER:O	1:M:1177:HIS:ND1	2.34	0.54
1:E:1182:ARG:HH21	1:E:1184:HIS:CE1	2.26	0.53
1:E:196:ARG:HE	1:E:1150:ARG:HH12	1.57	0.53
2:J:18:DG:H4'	2:J:19:DG:OP1	2.08	0.53
1:M:1182:ARG:HG2	1:M:1205:VAL:HG12	1.90	0.53
3:O:9:DT:H2"	3:O:10:DT:H5"	1.91	0.53
1:E:1134:ARG:NH2	1:E:1166:ASP:O	2.28	0.53
1:A:1182:ARG:HH21	1:A:1184:HIS:HE1	1.56	0.53
1:I:1113:ARG:O	1:I:1124:THR:N	2.40	0.53
1:M:1105:ASP:HB3	1:M:1135:LEU:HD22	1.90	0.53
1:E:1106:ILE:HG22	1:E:1179:TYR:HB3	1.91	0.52
1:M:1225:LEU:O	1:M:1229:ASN:ND2	2.33	0.51
3:K:24:DA:H2"	3:K:25:DG:O5'	2.10	0.51
1:M:1076:ILE:HD12	1:M:1162:ASN:HB3	1.93	0.51
1:E:1221:LYS:O	1:E:1224:GLN:HG2	2.11	0.50
1:I:1075:ILE:HD13	1:I:1214:VAL:HG23	1.93	0.49
1:A:1182:ARG:HG2	1:A:1205:VAL:HG22	1.95	0.49
1:A:1181:PRO:HG3	1:A:1211:PHE:HE2	1.77	0.49
1:I:1116:PHE:CD2	1:I:1218:GLN:HB3	2.48	0.49
1:I:1225:LEU:O	1:I:1229:ASN:ND2	2.31	0.49
3:C:8:DC:H2'	3:C:9:DT:C6	2.47	0.49
1:A:1135:LEU:HD12	1:E:1135:LEU:HD12	1.94	0.49
2:N:6:DC:H2"	2:N:7:DA:C8	2.47	0.49
1:M:1134:ARG:NH2	1:M:1166:ASP:O	2.31	0.49
1:M:1114:TYR:HE2	1:M:1178:LYS:HD2	1.78	0.49
1:A:1101:ILE:HG12	1:A:1142:PRO:HB3	1.95	0.48
1:E:177:THR:H	1:E:180:GLN:NE2	2.11	0.48
1:I:1082:ARG:HD2	3:K:16:DG:H5"	1.94	0.48
1:A:1135:LEU:HD11	1:E:1103:LEU:HD22	1.96	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:1063:LEU:HD12	1:I:1066:LYS:HD3	1.95	0.48
2:N:19:DG:H1	3:O:5:DC:H42	1.60	0.48
1:A:1075:ILE:HD13	1:A:1214:VAL:HG23	1.96	0.48
2:F:18:DG:H2"	2:F:19:DG:O5'	2.14	0.48
1:I:1235:GLY:HA2	2:J:9:DC:H5'	1.95	0.48
1:E:196:ARG:NE	1:E:1150:ARG:HH12	2.11	0.47
3:G:7:DA:H2"	3:G:8:DC:C6	2.48	0.47
3:G:24:DA:H2"	3:G:25:DG:O5'	2.14	0.47
1:I:1182:ARG:HH21	1:I:1184:HIS:HE1	1.60	0.47
1:I:160:ARG:NH2	1:I:196:ARG:HH22	2.13	0.47
1:E:189:ARG:O	1:E:193:LYS:HG2	2.15	0.47
2:N:10:DT:H1'	2:N:11:DT:H5'	1.95	0.47
1:I:1182:ARG:HG2	1:I:1205:VAL:HG22	1.97	0.47
1:I:161:TYR:HD2	3:K:5:DC:OP1	1.98	0.46
1:E:1227:ILE:O	1:E:1237:ARG:NE	2.45	0.46
1:E:1076:ILE:HG22	1:E:1160:LEU:HB3	1.98	0.45
1:E:1112:HIS:HB3	1:E:1123:VAL:HG13	1.99	0.45
1:I:1116:PHE:CZ	1:I:1119:ASN:HA	2.50	0.45
1:A:185:PHE:O	1:A:189:ARG:HG3	2.16	0.45
1:A:1095:ASN:HB3	1:A:1098:THR:HB	1.98	0.45
1:I:1106:ILE:HG22	1:I:1179:TYR:HB3	1.98	0.45
3:O:24:DA:H2"	3:O:25:DG:O5'	2.17	0.45
1:A:1106:ILE:HG22	1:A:1179:TYR:HB3	1.99	0.44
1:I:1110:ASP:OD2	1:I:1114:TYR:OH	2.26	0.44
1:M:1181:PRO:HG3	1:M:1211:PHE:HE1	1.82	0.44
2:B:10:DT:H1'	2:B:11:DT:H5'	1.98	0.44
2:J:10:DT:H1'	2:J:11:DT:H5'	1.99	0.44
1:A:1107:VAL:HG11	1:A:1182:ARG:NH1	2.31	0.44
1:M:161:TYR:HD2	3:O:5:DC:OP1	2.01	0.44
1:E:1081:ARG:NH2	3:G:18:DG:OP2	2.45	0.44
1:I:1175:SER:O	1:I:1177:HIS:ND1	2.39	0.44
1:M:1077:THR:HG23	1:M:1081:ARG:HB2	2.00	0.44
1:I:1176:MET:N	1:I:1213:ALA:O	2.48	0.44
1:A:1208:GLU:OE1	1:A:1208:GLU:N	2.47	0.44
1:M:1063:LEU:HD12	1:M:1066:LYS:HE3	1.99	0.44
1:A:1076:ILE:HD13	1:A:1213:ALA:HB1	2.00	0.43
1:A:1105:ASP:OD1	1:A:1184:HIS:NE2	2.48	0.43
2:F:10:DT:H1'	2:F:11:DT:H5'	2.00	0.43
2:B:6:DC:H2"	2:B:7:DA:C8	2.54	0.43
1:M:1208:GLU:OE1	1:M:1208:GLU:N	2.44	0.43
1:A:1062:GLU:O	1:A:1066:LYS:HG3	2.19	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:1095:ASN:HB3	1:E:1098:THR:HB	1.99	0.43
1:A:1112:HIS:HD2	1:A:1123:VAL:HG11	1.84	0.43
2:F:18:DG:H1	3:G:6:DC:H42	1.67	0.43
1:A:189:ARG:NH1	3:C:6:DC:OP2	2.48	0.43
1:M:189:ARG:HD2	3:O:6:DC:OP2	2.19	0.43
3:K:16:DG:H2"	3:K:17:DT:H5'	2.01	0.42
1:E:1062:GLU:CD	1:E:1062:GLU:H	2.23	0.42
1:I:1208:GLU:OE1	1:I:1208:GLU:N	2.47	0.42
2:B:13:DG:H2"	2:B:14:DA:C8	2.54	0.42
1:E:1065:LEU:O	1:E:1069:GLU:HG2	2.20	0.42
2:N:13:DG:H2"	2:N:14:DA:C8	2.54	0.42
1:M:1116:PHE:CD2	1:M:1218:GLN:HB3	2.54	0.42
1:M:1182:ARG:HH21	1:M:1184:HIS:CE1	2.37	0.42
1:A:1084:PHE:O	1:A:1226:LYS:NZ	2.37	0.42
3:C:17:DT:H2"	3:C:18:DG:C8	2.54	0.42
1:I:1110:ASP:N	1:I:1110:ASP:OD1	2.53	0.42
1:M:1105:ASP:OD1	1:M:1184:HIS:NE2	2.49	0.42
3:C:5:DC:H2'	3:C:6:DC:C6	2.55	0.42
1:E:1058:LEU:HB2	1:E:1206:PHE:CE2	2.54	0.42
3:K:5:DC:H2'	3:K:6:DC:C6	2.55	0.42
3:O:8:DC:H2'	3:O:9:DT:C6	2.55	0.41
1:A:1110:ASP:N	1:A:1110:ASP:OD1	2.53	0.41
1:A:1182:ARG:HH21	1:A:1184:HIS:CE1	2.37	0.41
1:E:1062:GLU:O	1:E:1066:LYS:HG3	2.20	0.41
1:I:185:PHE:O	1:I:189:ARG:HG3	2.20	0.41
1:M:1105:ASP:OD1	1:M:1182:ARG:HB2	2.19	0.41
1:I:140:PRO:N	1:I:141:ARG:HA	2.35	0.41
1:M:1065:LEU:O	1:M:1069:GLU:HG2	2.20	0.41
1:I:1077:THR:HG23	1:I:1081:ARG:HB2	2.03	0.41
1:M:1110:ASP:OD1	1:M:1110:ASP:N	2.53	0.41
1:M:1166:ASP:HA	1:M:1167:PRO:HD3	1.88	0.41
2:N:19:DG:N2	3:O:5:DC:N3	2.63	0.41
1:I:1073:GLU:HB2	1:I:1222:ILE:HG21	2.02	0.41
1:I:1106:ILE:CG2	1:I:1179:TYR:HB3	2.51	0.41
1:M:1082:ARG:HD2	3:O:16:DG:H5"	2.03	0.41
1:M:1091:VAL:HG13	1:M:1149:MET:HG2	2.03	0.41
1:M:1101:ILE:HG12	1:M:1142:PRO:HB3	2.02	0.41
1:E:1164:HIS:HA	1:E:1172:ILE:HD13	2.03	0.41
3:G:16:DG:H2"	3:G:17:DT:H5'	2.03	0.41
3:O:16:DG:H2"	3:O:17:DT:H5'	2.03	0.40
1:E:1099:LYS:HG2	1:E:1144:THR:HG22	2.02	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:1110:ASP:OD1	1:E:1110:ASP:N	2.53	0.40
1:M:1186:VAL:HG12	1:M:1187:LYS:H	1.87	0.40
1:A:1076:ILE:HD12	1:A:1162:ASN:HB3	2.03	0.40

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	220/285 (77%)	214 (97%)	6 (3%)	0	100	100
1	E	221/285 (78%)	215 (97%)	6 (3%)	0	100	100
1	I	218/285 (76%)	209 (96%)	9 (4%)	0	100	100
1	M	223/285 (78%)	215 (96%)	8 (4%)	0	100	100
All	All	882/1140 (77%)	853 (97%)	29 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	202/253 (80%)	201 (100%)	1 (0%)	88	96
1	E	205/253 (81%)	202 (98%)	3 (2%)	65	87

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	204/253 (81%)	201 (98%)	3 (2%)	65	87
1	M	205/253 (81%)	203 (99%)	2 (1%)	76	91
All	All	816/1012 (81%)	807 (99%)	9 (1%)	73	90

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1059	HIS
1	E	146	GLN
1	E	1061	ARG
1	E	1062	GLU
1	I	141	ARG
1	I	143	LEU
1	I	1204	HIS
1	M	1150	ARG
1	M	1220	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	180	GLN
1	E	1156	GLN
1	I	1156	GLN

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 ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

2 ligands 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	$\#|Z| > 2$
4	PG4	C	2025	-	12,12,12	0.68	0	11,11,11	1.50	0
4	PG4	A	2239	-	12,12,12	0.67	0	11,11,11	1.52	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
4	PG4	C	2025	-	-	5/10/10/10	-
4	PG4	A	2239	-	-	6/10/10/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	2025	PG4	O4-C7-C8-O5
4	C	2025	PG4	O3-C5-C6-O4
4	A	2239	PG4	C6-C5-O3-C4
4	C	2025	PG4	C8-C7-O4-C6
4	A	2239	PG4	O3-C5-C6-O4
4	A	2239	PG4	O2-C3-C4-O3
4	A	2239	PG4	C1-C2-O2-C3
4	A	2239	PG4	C5-C6-O4-C7
4	C	2025	PG4	O1-C1-C2-O2
4	A	2239	PG4	O1-C1-C2-O2
4	C	2025	PG4	O2-C3-C4-O3

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 ⓘ

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	226/285 (79%)	-0.01	3 (1%)	77	51	41, 63, 84, 105	0
1	E	227/285 (79%)	0.08	1 (0%)	92	79	44, 66, 98, 117	0
1	I	226/285 (79%)	0.26	4 (1%)	68	40	47, 79, 103, 119	0
1	M	229/285 (80%)	0.15	4 (1%)	70	41	49, 74, 99, 117	0
2	B	21/22 (95%)	-0.36	0	100	100	68, 77, 92, 101	0
2	F	21/22 (95%)	-0.31	0	100	100	65, 77, 88, 108	0
2	J	21/22 (95%)	-0.25	0	100	100	70, 81, 92, 120	0
2	N	21/22 (95%)	-0.32	0	100	100	48, 70, 85, 94	0
3	C	20/22 (90%)	-0.23	0	100	100	64, 86, 98, 101	0
3	G	21/22 (95%)	-0.28	0	100	100	65, 76, 102, 111	0
3	K	22/22 (100%)	-0.12	0	100	100	67, 80, 108, 116	0
3	O	22/22 (100%)	-0.16	0	100	100	63, 69, 98, 117	0
All	All	1077/1316 (81%)	0.06	12 (1%)	80	56	41, 72, 100, 120	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	M	1201	PHE	3.2
1	A	1118	ASP	3.1
1	A	1128	GLU	2.8
1	I	172	SER	2.8
1	I	1062	GLU	2.7
1	M	192	SER	2.7
1	I	1119	ASN	2.5
1	A	1130	ALA	2.5
1	M	1202	SER	2.5
1	I	1091	VAL	2.4
1	M	1220	HIS	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	197	GLN	2.1

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 ⓘ

There are no monosaccharides in this entry.

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
4	PG4	C	2025	13/13	0.78	0.28	75,91,106,107	0
4	PG4	A	2239	13/13	0.83	0.23	61,68,83,85	0

6.5 Other polymers ⓘ

There are no such residues in this entry.