Full wwPDB X-ray Structure Validation Report (i)

Aug 8, 2020 – 03:06 PM BST

PDB ID	:	5CJS
Title	:	Crystal structure of a monomeric influenza hemagglutinin stem in complex with an broadly neutralizing antibody CR9114
Authors	:	Zhu, X.; Wilson, I.A.
Deposited on	:	2015-07-15
Resolution	:	4.30 Å(reported)

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

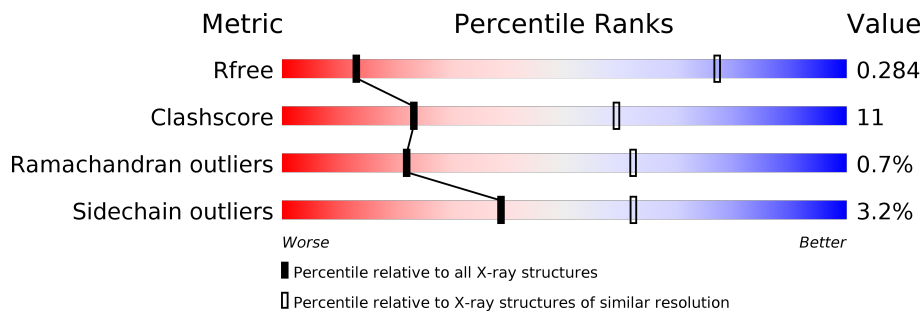|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 4.30 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		0.284
Clashscore		11
Ramachandran outliers		0.7%
Sidechain outliers		3.2%

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	1014 (4.80-3.80)
Clashscore	141614	1077 (4.80-3.80)
Ramachandran outliers	138981	1029 (4.80-3.80)
Sidechain outliers	138945	1012 (4.80-3.80)

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 on 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%

Mol	Chain	Length	Quality of chain
1	E	215	77% 18% • •
1	L	215	77% 20% • •
2	F	230	73% 21% 6%
2	H	230	73% 18% • • 6%
3	C	62	44% 37% • 18%
3	J	62	56% 19% 8% 16%
4	D	191	42% 29% • • 25%

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	K	191	59%　　19%　　18%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 9576 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 CR9114 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	211	Total 1568	C 978	N 266	O 320	S 4	0	0	0
1	E	211	Total 1568	C 978	N 266	O 320	S 4	0	0	0

- Molecule 2 is a protein called CR9114 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	217	Total 1613	C 1017	N 269	O 320	S 7	0	0	0
2	F	217	Total 1613	C 1017	N 269	O 320	S 7	0	0	0

- Molecule 3 is a protein called Designed influenza hemagglutinin stem #4454, HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	51	Total 387	C 235	N 70	O 79	S 3	0	0	0
3	J	52	Total 397	C 242	N 72	O 80	S 3	0	0	0

- Molecule 4 is a protein called Designed influenza hemagglutinin stem #4454, HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	143	Total 1154	C 717	N 195	O 235	S 7	0	0	0
4	K	156	Total 1234	C 766	N 209	O 252	S 7	0	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	C	1	Total	C	N	O	0	0
			14	8	1	5		
5	J	1	Total	C	N	O	0	0
			14	8	1	5		
5	K	1	Total	C	N	O	0	0
			14	8	1	5		

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 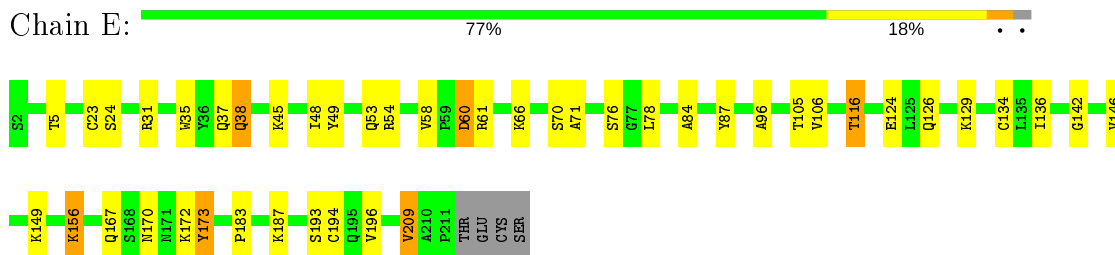quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
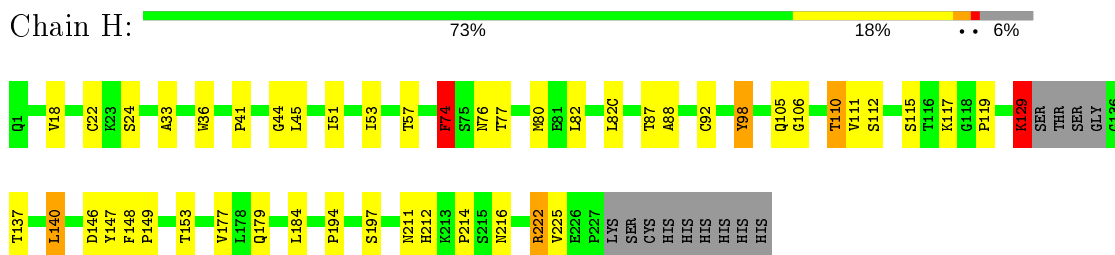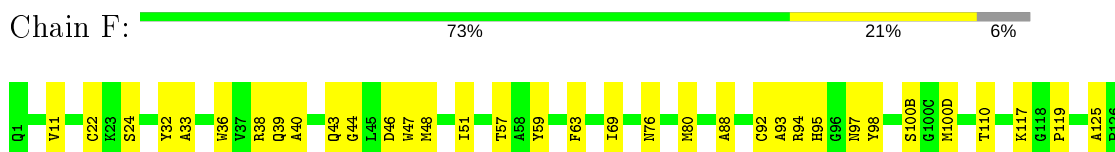
- Molecule 1: CR9114 light chain

Chain L: 77% 20%

- Molecule 1: CR9114 light chain

Chain E: 77% 18%

- Molecule 2: CR9114 heavy chain

Chain H: 73% 18% 6%

- Molecule 2: CR9114 heavy chain

Chain F: 73% 21% 6%

S127 S128 K129 THR SER GLY G136 T137 A138 A139 L140 D146 P151 V152 T153 L166 V191 N211 H212 K213 P214 T217 V225 E226 P227 LYS SER CYS HIS HIS HIS HIS HIS

• Molecule 3: Designed influenza hemagglutinin stem #4454, HA1

Chain C: 44% 37% · 18%

D11 T12 I13 C14 I15 G16 Y17 H18 A19 V26 V34 T35 V36 T37 H38 S39 L43 E44 N45 GLY GLY GLY LYS TYR VAL C53 S54 A55 K56 L57 R58 M59 V60 R64 N65 K66 P67 S68 LYS GLN SER GLN

• Molecule 3: Designed influenza hemagglutinin stem #4454, HA1

Chain J: 56% 19% 8% 16%

D11 G16 V26 V29 L30 T35 V36 E44 ASN GLY GLY GLY LYS TYR VAL C53 S54 A55 K56 L57 V60 L63 K66 P67 S68 K69 Q70 SER GLN

• Molecule 4: Designed influenza hemagglutinin stem #4454, HA2

Chain D: 42% 29% · · 25%

GLY LEU PHE GLY ALA I6 A7 G8 E11 G12 G13 W14 T15 G16 D19 G20 W21 Y24 H25 H26 Q27 S32 G33 Y34 Q42 A44 I45 I48 T49 N60 THR GLN TYR THR ALA ILE GLY LYS GLU TYR ASN LYS SER GLU ARG MET LYS GLN ILE GLU ASP LYS

ILE E84 E85 I86 E87 S88 K89 Q90 I91 W92 N95 A96 E97 L98 L99 V100 L101 L102 R106 T107 L108 D109 N117 L118 Y119 E120 K121 V122 K123 S124 Q125 L126 K131 E132 I133 G134 N135 F138 E139 H142 K143 C144 N145 D146 E147 C148 V152 K153 D158 Y159 P160 K161

Y162 E171 LYS ILE ASP GLY ARG SER LEU VAL PRO ARG GLY SER PRO GLY HIS HIS HIS HIS HIS HIS

• Molecule 4: Designed influenza hemagglutinin stem #4454, HA2

Chain K: 59% 19% · 18%

GLY LEU PHE G4 A5 I6 A7 G8 F9 T10 G13 W14 M17 V18 D19 G20 W21 Y22 G23 A36 D37 Q38 T41 Q42 N43 A44 I45 N46 G47 I48 M59 N60 T61 Q62 Y63 T64 ALA ILE GLY LYS GLU TYR ASN LYS SER GLU ARG MET LYS GLN ILE GLU ASP LYS I83

N92 C93 Y94 N95 A96 E97 L102 H111 V115 K116 Y119 E120 K123 N128 G136 C137 F138 E164 L168 N169 R170 D174 G175 R176 S177 LEU VAL PRO ARG GLY SER PRO GLY HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	110.87Å 110.87Å 359.16Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.01 – 4.30 48.01 – 4.29	Depositor EDS
% Data completeness (in resolution range)	99.5 (48.01-4.30) 99.6 (48.01-4.29)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.21	Depositor
$< I/\sigma(I) >$ [1]	1.96 (at 4.29Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.262 , 0.303 0.277 , 0.284	Depositor DCC
R_{free} test set	933 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	147.6	Xtriage
Anisotropy	0.615	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 165.6	EDS
L-test for twinning[2]	$< |L| > = 0.43, < L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.069 for -h,-k,l	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	9576	wwPDB-VP
Average B, all atoms (Å^2)	179.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.15% 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: 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	E	0.34	0/1606	0.71	2/2193 (0.1%)
1	L	0.39	0/1606	0.85	6/2193 (0.3%)
2	F	0.38	0/1652	0.69	0/2251
2	H	0.47	1/1652 (0.1%)	0.81	5/2251 (0.2%)
3	C	0.51	0/389	1.08	2/527 (0.4%)
3	J	0.47	0/399	0.98	2/539 (0.4%)
4	D	0.61	2/1176 (0.2%)	1.08	10/1581 (0.6%)
4	K	0.50	0/1257	0.92	3/1692 (0.2%)
All	All	0.45	3/9737 (0.0%)	0.85	30/13227 (0.2%)

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
4	D	0	1
4	K	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	129	LYS	CE-NZ	8.88	1.71	1.49
4	D	143	LYS	CE-NZ	6.59	1.65	1.49
4	D	11	GLU	CG-CD	5.08	1.59	1.51

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	190	ARG	NE-CZ-NH2	-11.53	114.53	120.30
4	K	83	ILE	CG1-CB-CG2	11.24	136.12	111.40
1	L	19	VAL	CG1-CB-CG2	8.04	123.76	110.90
4	D	131	LYS	CB-CA-C	-7.86	94.68	110.40
1	E	209	VAL	CG1-CB-CG2	7.55	122.97	110.90
1	L	156	LYS	CA-CB-CG	7.23	129.31	113.40
3	J	56	LYS	N-CA-CB	-6.92	98.15	110.60
4	D	131	LYS	CA-CB-CG	6.76	128.28	113.40
4	D	143	LYS	CD-CE-NZ	-6.52	96.70	111.70
1	L	155	VAL	CG1-CB-CG2	6.27	120.94	110.90
4	D	131	LYS	N-CA-CB	6.21	121.77	110.60
4	D	134	GLY	N-CA-C	-6.06	97.96	113.10
4	D	143	LYS	CG-CD-CE	6.05	130.05	111.90
2	H	222	ARG	CG-CD-NE	6.00	124.40	111.80
2	H	74	PHE	CB-CG-CD2	-6.00	116.60	120.80
2	H	74	PHE	CB-CG-CD1	5.98	124.99	120.80
4	D	119	TYR	CB-CA-C	-5.98	98.44	110.40
2	H	129	LYS	CB-CG-CD	5.87	126.87	111.60
1	L	190	ARG	NE-CZ-NH1	5.57	123.08	120.30
1	E	156	LYS	CB-CG-CD	-5.52	97.25	111.60
4	D	89	LYS	CA-CB-CG	5.50	125.50	113.40
3	C	66	LYS	CA-CB-CG	5.45	125.38	113.40
1	L	19	VAL	CA-CB-CG2	5.39	118.98	110.90
4	D	11	GLU	CA-CB-CG	5.19	124.83	113.40
4	K	123	LYS	CD-CE-NZ	-5.18	99.78	111.70
3	C	66	LYS	CD-CE-NZ	-5.17	99.81	111.70
3	J	35	THR	OG1-CB-CG2	5.14	121.83	110.00
4	K	116	LYS	CG-CD-CE	-5.13	96.51	111.90
4	D	119	TYR	CB-CG-CD2	-5.13	117.92	121.00
2	H	110	THR	OG1-CB-CG2	5.04	121.60	110.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	D	159	TYR	Sidechain
4	K	59	MET	Peptide

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	E	1568	0	1521	29	1
1	L	1568	0	1521	26	0
2	F	1613	0	1572	32	0
2	H	1613	0	1572	34	1
3	C	387	0	386	20	0
3	J	397	0	401	11	0
4	D	1154	0	1073	52	0
4	K	1234	0	1134	33	0
5	C	14	0	13	0	0
5	J	14	0	13	0	0
5	K	14	0	13	0	0
All	All	9576	0	9219	212	1

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 11.

All (212) 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:H:129:LYS:CE	2:H:129:LYS:NZ	1.71	1.52
4:D:133:ILE:HD11	4:D:139:GLU:HB2	1.45	0.94
3:J:26:VAL:HG21	3:J:60:VAL:HG22	1.54	0.87
1:E:142:GLY:HA3	1:E:173:TYR:CD2	2.17	0.80
3:C:26:VAL:HG21	3:C:60:VAL:HG22	1.65	0.78
2:H:117:LYS:NZ	2:H:146:ASP:O	2.17	0.78
4:D:159:TYR:HB3	4:D:160:PRO:HD3	1.66	0.77
1:L:142:GLY:HA3	1:L:173:TYR:CD2	2.22	0.75
1:E:116:THR:HG23	2:F:128:SER:HB2	1.68	0.75
4:K:17:MET:SD	4:K:23:GLY:HA3	2.27	0.74
4:D:123:LYS:HB2	4:D:138:PHE:HZ	1.52	0.73
4:K:9:PHE:HD1	4:K:10:THR:N	1.87	0.72
3:C:39:SER:OG	3:C:58:ARG:NH2	2.21	0.72
2:F:117:LYS:NZ	2:F:146:ASP:O	2.22	0.70
4:D:119:TYR:CE2	4:D:138:PHE:HE1	2.09	0.70
4:D:119:TYR:CE2	4:D:138:PHE:CE1	2.81	0.69
1:E:60:ASP:OD1	1:E:60:ASP:N	2.28	0.66
3:J:66:LYS:O	4:K:13:GLY:N	2.26	0.65
2:H:24:SER:HB3	2:H:76:ASN:HB3	1.77	0.65
4:D:133:ILE:HD11	4:D:139:GLU:CB	2.22	0.64

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:L:5:THR:HB	1:L:24:SER:HB2	1.80	0.64
4:K:62:GLN:HG3	4:K:63:TYR:N	2.14	0.63
2:F:166:LEU:HD21	2:F:191:VAL:HG21	1.80	0.63
3:C:17:TYR:OH	3:C:66:LYS:HB3	1.99	0.63
3:C:53:CYS:SG	3:C:54:SER:N	2.74	0.61
3:C:14:CYS:N	4:D:25:HIS:O	2.32	0.60
2:H:179:GLN:HG3	2:H:184:LEU:O	2.00	0.60
4:D:123:LYS:HB2	4:D:138:PHE:CZ	2.36	0.60
2:F:140:LEU:HD13	2:F:225:VAL:HG11	1.84	0.60
1:L:38:GLN:O	1:L:84:ALA:HB1	2.02	0.60
2:H:112:SER:HB3	2:H:148:PHE:HZ	1.65	0.60
4:D:126:LEU:HD12	4:D:138:PHE:CD2	2.37	0.59
1:L:108:GLN:HB2	1:L:140:TYR:CD1	2.38	0.59
1:L:108:GLN:HB2	1:L:140:TYR:CE1	2.38	0.59
1:L:32:SER:HB3	1:L:50:SER:HA	1.85	0.59
1:E:136:ILE:HG12	1:E:196:VAL:HG21	1.86	0.58
2:F:119:PRO:HD2	2:F:217:THR:HG21	1.84	0.58
1:L:145:THR:OG1	1:L:197:THR:HB	2.04	0.58
3:C:11:ASP:OD2	4:D:144:CYS:N	2.29	0.58
3:C:14:CYS:HB2	4:D:25:HIS:HB3	1.85	0.57
4:K:7:ALA:HB3	4:K:14:TRP:HE1	1.69	0.57
2:H:36:TRP:CE2	2:H:80:MET:HB2	2.38	0.57
3:C:26:VAL:HG21	3:C:60:VAL:CG2	2.34	0.57
2:H:129:LYS:HB2	2:H:137:THR:O	2.05	0.57
3:J:69:LYS:HG3	3:J:70:GLN:HG3	1.87	0.56
4:K:92:TRP:O	4:K:95:ASN:HB2	2.04	0.56
1:L:83:GLU:OE1	1:L:106:VAL:N	2.37	0.56
3:J:53:CYS:SG	3:J:54:SER:N	2.79	0.56
4:D:95:ASN:OD1	4:K:95:ASN:OD1	2.23	0.56
1:L:106(A):LEU:HA	1:L:140:TYR:OH	2.06	0.56
4:D:106:ARG:NH1	3:J:29:VAL:O	2.39	0.56
1:L:134:CYS:HG	1:L:194:CYS:HG	1.53	0.55
2:H:98:TYR:N	4:D:42:GLN:OE1	2.36	0.55
3:J:36:VAL:HG21	3:J:60:VAL:HG13	1.89	0.55
2:H:33:ALA:HB2	2:H:98:TYR:O	2.07	0.55
1:E:149:LYS:HB2	1:E:193:SER:HB2	1.89	0.55
2:F:51:ILE:HD12	2:F:57:THR:HG22	1.88	0.55
4:D:16:GLY:HA3	4:D:34:TYR:CE2	2.42	0.55
2:H:140:LEU:HD13	2:H:225:VAL:HG11	1.89	0.55
3:C:34:VAL:CG1	3:C:64:ARG:HG3	2.37	0.54
4:D:8:GLY:H	4:D:14:TRP:HE1	1.53	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:161:LYS:HG2	4:D:162:TYR:CE1	2.42	0.54
4:D:92:TRP:O	4:D:95:ASN:HB2	2.08	0.54
3:C:19:ALA:HB2	4:D:13:GLY:HA3	1.90	0.54
2:H:53:ILE:HG13	4:D:49:THR:OG1	2.08	0.54
1:E:183:PRO:O	1:E:187:LYS:HG3	2.08	0.53
3:C:60:VAL:HG11	4:D:108:LEU:HD21	1.90	0.53
4:K:128:ASN:ND2	4:K:170:ARG:HH12	2.07	0.53
3:C:36:VAL:HG21	3:C:60:VAL:HG13	1.90	0.53
4:D:134:GLY:O	4:D:135:ASN:OD1	2.27	0.52
2:H:146:ASP:HB3	2:H:184:LEU:HD13	1.89	0.52
2:F:24:SER:HB3	2:F:76:ASN:HB3	1.91	0.52
1:E:134:CYS:HG	1:E:194:CYS:HG	1.58	0.52
1:L:115:VAL:O	1:L:207:LYS:HE2	2.11	0.51
4:D:86:ILE:HA	4:D:89:LYS:HE2	1.93	0.51
2:H:112:SER:HB3	2:H:148:PHE:CZ	2.44	0.51
1:L:80:SER:HA	1:L:106:VAL:HG11	1.92	0.51
1:E:78:LEU:HD21	1:E:106:VAL:HG22	1.93	0.51
2:F:129:LYS:HD2	2:F:138:ALA:HB2	1.93	0.51
1:E:96:ALA:HB3	2:F:47:TRP:CG	2.46	0.51
4:K:119:TYR:CE2	4:K:136:GLY:HA2	2.46	0.51
3:C:68:SER:OG	3:C:68:SER:O	2.27	0.50
4:D:24:TYR:HB2	4:D:153:LYS:HE2	1.92	0.50
2:F:128:SER:O	2:F:129:LYS:HD3	2.11	0.50
4:K:128:ASN:O	4:K:170:ARG:NH2	2.44	0.50
1:E:37:GLN:O	1:E:45:LYS:N	2.42	0.50
4:K:119:TYR:CE1	4:K:138:PHE:CZ	3.00	0.50
1:E:38:GLN:O	1:E:84:ALA:HB1	2.12	0.49
2:H:51:ILE:HD12	2:H:57:THR:HG22	1.94	0.49
3:C:66:LYS:O	4:D:13:GLY:N	2.29	0.49
4:D:102:LEU:O	4:D:106:ARG:HG3	2.13	0.49
2:F:11:VAL:HG22	2:F:110:THR:HB	1.94	0.49
1:L:136:ILE:HG12	1:L:196:VAL:HG21	1.94	0.49
4:D:161:LYS:HB3	4:D:162:TYR:HD1	1.78	0.49
4:D:19:ASP:OD1	4:D:19:ASP:N	2.44	0.48
2:F:51:ILE:CD1	2:F:57:THR:HG22	2.42	0.48
1:E:49:TYR:CZ	1:E:53:GLN:HB3	2.48	0.48
4:K:164:GLU:O	4:K:168:LEU:HG	2.13	0.48
4:D:161:LYS:HB3	4:D:162:TYR:CD1	2.49	0.48
4:D:8:GLY:N	4:D:14:TRP:HE1	2.12	0.48
4:D:98:LEU:HD22	4:K:102:LEU:HD11	1.96	0.48
2:F:212:HIS:CD2	2:F:214:PRO:HD2	2.49	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:18:HIS:HB2	4:D:20:GLY:O	2.13	0.48
4:D:148:CYS:HB2	4:D:162:TYR:OH	2.14	0.48
4:K:9:PHE:HD2	4:K:115:VAL:HG12	1.78	0.48
4:D:88:SER:O	4:D:91:ILE:HB	2.14	0.48
2:F:93:ALA:HB1	2:F:100(D):MET:HB3	1.96	0.48
1:E:49:TYR:O	1:E:53:GLN:HB2	2.14	0.47
2:H:87:THR:OG1	2:H:111:VAL:N	2.31	0.47
2:H:74:PHE:H	2:H:74:PHE:HD1	1.62	0.47
2:H:36:TRP:CD2	2:H:80:MET:HB2	2.50	0.47
2:H:146:ASP:HA	2:H:184:LEU:HB3	1.97	0.47
1:L:37:GLN:HB2	1:L:47:LEU:HD13	1.96	0.47
4:D:162:TYR:N	4:D:162:TYR:HD1	2.13	0.47
2:F:59:TYR:HE1	2:F:69:ILE:HG13	1.80	0.47
1:L:35:TRP:HB2	1:L:48:ILE:HB	1.97	0.46
2:F:95:HIS:HB3	2:F:100(B):SER:OG	2.15	0.46
3:C:56:LYS:O	3:C:57:LEU:HD23	2.14	0.46
4:K:61:THR:OG1	4:K:92:TRP:NE1	2.40	0.46
1:E:35:TRP:HB2	1:E:48:ILE:HB	1.98	0.46
2:F:129:LYS:HB3	2:F:137:THR:O	2.15	0.46
2:H:22:CYS:O	2:H:77:THR:HA	2.15	0.46
1:E:66:LYS:HA	1:E:71:ALA:HA	1.98	0.46
4:K:9:PHE:CD1	4:K:10:THR:N	2.76	0.46
4:K:20:GLY:HA3	4:K:36:ALA:HB1	1.98	0.46
4:K:42:GLN:NE2	4:K:46:ASN:OD1	2.48	0.46
4:D:97:GLU:O	4:D:101:LEU:HG	2.16	0.46
2:H:129:LYS:NZ	2:H:129:LYS:CD	2.70	0.46
1:E:146:VAL:HG22	1:E:196:VAL:HG22	1.96	0.46
1:E:38:GLN:HE22	2:F:39:GLN:HE22	1.64	0.46
4:D:145:ASN:ND2	4:D:147:GLU:HG2	2.31	0.45
4:D:21:TRP:CZ3	4:D:45:ILE:HG13	2.51	0.45
1:E:87:TYR:CE1	2:F:44:GLY:HA2	2.50	0.45
4:K:19:ASP:N	4:K:19:ASP:OD1	2.46	0.45
1:L:78:LEU:HD21	1:L:106:VAL:HG22	1.98	0.45
1:L:150:ALA:N	1:L:153:SER:O	2.44	0.45
4:D:44:ALA:O	4:D:48:ILE:HG12	2.16	0.45
2:F:125:ALA:O	2:F:127:SER:N	2.47	0.45
2:F:39:GLN:O	2:F:88:ALA:HB1	2.17	0.45
4:D:27:GLN:HG3	4:D:32:SER:HB3	1.98	0.45
4:D:86:ILE:HA	4:D:89:LYS:CE	2.46	0.45
3:C:35:THR:HG22	3:C:65:ASN:HB3	1.98	0.45
2:F:97:ASN:ND2	4:K:38:GLN:OE1	2.49	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:159:TYR:HB3	4:D:160:PRO:CD	2.43	0.45
4:K:119:TYR:O	4:K:119:TYR:HD1	1.99	0.45
1:L:87:TYR:CE2	2:H:44:GLY:HA2	2.51	0.45
1:E:23:CYS:O	1:E:70:SER:HA	2.17	0.45
2:H:119:PRO:HB3	2:H:147:TYR:HB3	1.99	0.45
1:L:183:PRO:O	1:L:187:LYS:HG3	2.16	0.45
4:K:62:GLN:CG	4:K:63:TYR:N	2.71	0.44
4:D:7:ALA:HB1	4:D:8:GLY:HA3	1.99	0.44
1:L:192:TYR:O	1:L:208:THR:HG23	2.17	0.44
4:D:109:ASP:OD2	3:J:30:LEU:HD11	2.17	0.44
1:L:51:ASN:OD1	1:L:66:LYS:HD3	2.17	0.44
1:L:60:ASP:N	1:L:60:ASP:OD1	2.48	0.44
2:F:33:ALA:HB2	2:F:98:TYR:O	2.18	0.44
3:C:43:LEU:HD12	3:C:43:LEU:HA	1.78	0.44
4:D:142:HIS:ND1	4:D:162:TYR:CD2	2.85	0.44
1:E:61:ARG:HB2	1:E:76:SER:O	2.18	0.44
2:F:40:ALA:HB3	2:F:43:GLN:HG3	2.00	0.44
2:H:22:CYS:SG	2:H:92:CYS:SG	3.11	0.44
4:D:162:TYR:N	4:D:162:TYR:CD1	2.85	0.43
4:D:99:LEU:HD23	4:D:99:LEU:HA	1.79	0.43
1:E:124:GLU:HG2	1:E:129:LYS:O	2.18	0.43
3:J:63:LEU:HB3	4:K:111:HIS:CG	2.54	0.43
1:L:142:GLY:HA3	1:L:173:TYR:CG	2.53	0.43
2:H:148:PHE:HA	2:H:149:PRO:HA	1.81	0.43
2:H:194:PRO:HG2	2:H:197:SER:HB3	2.00	0.43
4:K:41:THR:HG22	4:K:45:ILE:HD12	2.00	0.43
1:L:186:TRP:HA	1:L:192:TYR:HE2	1.84	0.43
3:C:18:HIS:HD2	3:C:37:THR:HG21	1.83	0.43
2:F:153:THR:CG2	2:F:211:ASN:HB3	2.48	0.43
4:D:117:ASN:O	4:D:121:LYS:HB2	2.19	0.42
1:E:142:GLY:HA3	1:E:173:TYR:CE2	2.53	0.42
1:E:87:TYR:HE1	2:F:44:GLY:HA2	1.84	0.42
4:K:9:PHE:HD1	4:K:10:THR:H	1.66	0.42
1:E:31:ARG:HA	1:E:31:ARG:HD3	1.84	0.42
2:H:105:GLN:HG2	2:H:106:GLY:O	2.19	0.42
3:J:26:VAL:HG21	3:J:60:VAL:CG2	2.37	0.42
1:E:167:GLN:NE2	1:E:172:LYS:HB2	2.35	0.42
4:K:44:ALA:O	4:K:48:ILE:HG12	2.19	0.42
2:F:38:ARG:HG2	2:F:46:ASP:HB3	2.01	0.42
2:H:153:THR:CG2	2:H:211:ASN:HB3	2.50	0.42
2:H:18:VAL:HG12	2:H:82(C):LEU:HD11	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:45:LEU:HD23	2:H:45:LEU:HA	1.80	0.42
1:E:38:GLN:HE22	2:F:39:GLN:NE2	2.18	0.42
1:L:120:PRO:HB2	1:L:125:LEU:HD21	2.01	0.42
2:H:212:HIS:CD2	2:H:214:PRO:HD2	2.55	0.41
4:K:120:GLU:HA	4:K:123:LYS:HB3	2.02	0.41
4:D:142:HIS:CE1	4:D:162:TYR:CD2	3.07	0.41
2:H:222:ARG:HH11	2:H:222:ARG:HD3	1.69	0.41
2:H:41:PRO:HD3	2:H:88:ALA:HA	2.01	0.41
4:D:158:ASP:OD1	4:D:160:PRO:HD2	2.20	0.41
1:E:5:THR:HB	1:E:24:SER:HB2	2.02	0.41
2:H:82:LEU:HD12	2:H:82:LEU:HA	1.82	0.41
2:F:32:TYR:HB2	2:F:94:ARG:HD2	2.02	0.41
3:J:56:LYS:O	3:J:57:LEU:HD23	2.20	0.41
4:K:21:TRP:CZ3	4:K:45:ILE:HG13	2.55	0.41
4:K:7:ALA:HB3	4:K:14:TRP:NE1	2.34	0.41
3:C:15:ILE:CD1	4:D:122:VAL:HG21	2.50	0.41
4:D:125:GLN:NE2	4:D:152:VAL:O	2.42	0.41
4:K:111:HIS:O	4:K:115:VAL:HG23	2.21	0.41
1:E:167:GLN:NE2	1:E:170:ASN:OD1	2.51	0.41
2:F:36:TRP:CE2	2:F:80:MET:HB2	2.56	0.41
4:K:94:TYR:HA	4:K:97:GLU:HG2	2.03	0.41
1:L:23:CYS:O	1:L:70:SER:HA	2.21	0.41
2:H:74:PHE:CD1	2:H:74:PHE:N	2.89	0.41
3:J:16:GLY:HA3	4:K:14:TRP:CZ2	2.56	0.41
2:F:48:MET:HA	2:F:63:PHE:CD2	2.56	0.40
2:F:22:CYS:SG	2:F:92:CYS:SG	3.13	0.40
4:K:7:ALA:CB	4:K:14:TRP:HE1	2.34	0.40
1:E:54:ARG:NH1	1:E:58:VAL:O	2.48	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:115:SER:O	1:E:126:GLN:NE2[2_454]	2.05	0.15

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 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	E	209/215 (97%)	203 (97%)	6 (3%)	0	100	100
1	L	209/215 (97%)	203 (97%)	6 (3%)	0	100	100
2	F	213/230 (93%)	208 (98%)	4 (2%)	1 (0%)	29	68
2	H	213/230 (93%)	207 (97%)	5 (2%)	1 (0%)	29	68
3	C	47/62 (76%)	42 (89%)	5 (11%)	0	100	100
3	J	48/62 (77%)	45 (94%)	1 (2%)	2 (4%)	3	25
4	D	139/191 (73%)	138 (99%)	1 (1%)	0	100	100
4	K	152/191 (80%)	142 (93%)	6 (4%)	4 (3%)	5	35
All	All	1230/1396 (88%)	1188 (97%)	34 (3%)	8 (1%)	22	62

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	K	6	ILE
4	K	7	ALA
4	K	174	ASP
4	K	176	ARG
3	J	69	LYS
3	J	68	SER
2	H	216	ASN
2	F	151	PRO

5.3.2 Protein sidechains (i)

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	E	175/179 (98%)	168 (96%)	7 (4%)	31	56
1	L	175/179 (98%)	171 (98%)	4 (2%)	50	70
2	F	181/193 (94%)	180 (99%)	1 (1%)	86	92
2	H	181/193 (94%)	175 (97%)	6 (3%)	38	62
3	C	47/54 (87%)	45 (96%)	2 (4%)	29	55
3	J	48/54 (89%)	44 (92%)	4 (8%)	11	37
4	D	125/165 (76%)	119 (95%)	6 (5%)	25	52
4	K	130/165 (79%)	126 (97%)	4 (3%)	40	63
All	All	1062/1182 (90%)	1028 (97%)	34 (3%)	39	62

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	19	VAL
1	L	155	VAL
1	L	156	LYS
1	L	190	ARG
2	H	74	PHE
2	H	98	TYR
2	H	110	THR
2	H	129	LYS
2	H	140	LEU
2	H	177	VAL
3	C	12	THR
3	C	66	LYS
4	D	60	ASN
4	D	85	GLU
4	D	86	ILE
4	D	87	GLU
4	D	89	LYS
4	D	162	TYR
1	E	38	GLN
1	E	60	ASP
1	E	105	THR
1	E	116	THR
1	E	156	LYS
1	E	173	TYR
1	E	209	VAL
2	F	140	LEU
3	J	35	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	J	44	GLU
3	J	63	LEU
3	J	66	LYS
4	K	9	PHE
4	K	61	THR
4	K	83	ILE
4	K	119	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	L	189	HIS
4	D	95	ASN
1	E	38	GLN
2	F	216	ASN

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)

3 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
5	NAG	K	201	4	14,14,15	0.67	1 (7%)	17,19,21	0.82	1 (5%)
5	NAG	C	101	3	14,14,15	0.53	0	17,19,21	0.93	1 (5%)
5	NAG	J	501	3	14,14,15	0.40	0	17,19,21	0.65	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	K	201	4	-	0/6/23/26	0/1/1/1
5	NAG	C	101	3	-	0/6/23/26	0/1/1/1
5	NAG	J	501	3	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	K	201	NAG	C1-C2	2.21	1.55	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	101	NAG	C1-O5-C5	3.31	116.68	112.19
5	K	201	NAG	C1-O5-C5	2.65	115.79	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	J	501	NAG	O5-C5-C6-O6
5	J	501	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

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 (i)

6.1 Protein, DNA and RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.