Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 09:45 PM GMT

PDB ID : 4W5C

Title : Crystal structure analysis of cruzain with three Fragments: 1 (N-(1H-benzimi dazol-2-yl)-1,3-dimethyl-pyrazole-4-carboxamide), 6 (2-amino-4,6-difluoroben zothiazole) and 9 (N-(1H-benzimidazol-2-yl)-3-(4-fluorophenyl)-1H-pyrazole-4 -carboxamide).

Authors : Tochowicz, A.; McKerrow, J.H.; Craik, C.S.

Deposited on : 2014-08-17

Resolution : 3.27 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	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 3.27 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		0.249
Clashscore		8
Ramachandran outliers		0.9%
Sidechain outliers		7.4%
RSRZ outliers		3.5%

Worse .... *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
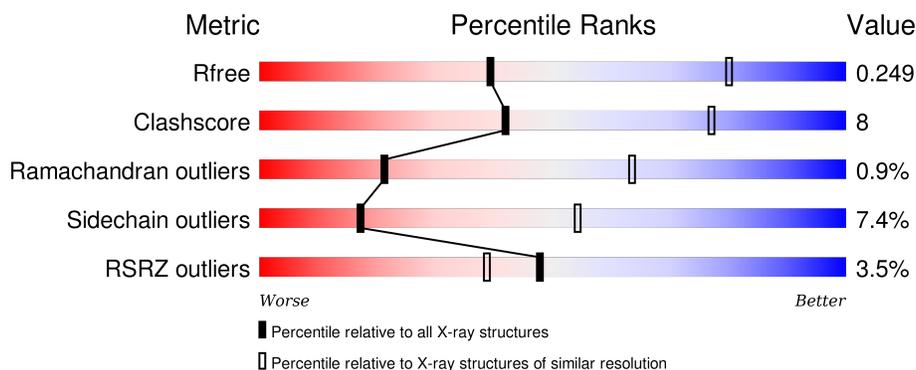
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	91344	1756 (3.36-3.20)
Clashscore	102246	1941 (3.36-3.20)
Ramachandran outliers	100387	1905 (3.36-3.20)
Sidechain outliers	100360	1903 (3.36-3.20)
RSRZ outliers	91569	1764 (3.36-3.20)

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. 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	216	5% / 81% / 17% / •
1	B	216	2% / 79% / 20%
1	C	216	2% / 77% / 21% / •
1	D	216	% / 76% / 22% / •
1	E	216	8% / 71% / 26% / •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	3H5	A	301	-	-	-	X
2	3H5	C	302	-	-	-	X
3	3H6	A	302	-	-	-	X
4	3H7	C	301	-	-	-	X
4	3H7	D	301	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8074 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 Cruzipain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	215	Total	C	N	O	S	0	0	0
			1592	995	266	320	11			
1	B	215	Total	C	N	O	S	0	0	0
			1592	995	266	320	11			
1	C	215	Total	C	N	O	S	0	0	0
			1592	995	266	320	11			
1	D	215	Total	C	N	O	S	0	0	0
			1592	995	266	320	11			
1	E	215	Total	C	N	O	S	0	0	0
			1588	992	265	320	11			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	25	SER	CYS	engineered mutation	UNP P25779
B	25	SER	CYS	engineered mutation	UNP P25779
C	25	SER	CYS	engineered mutation	UNP P25779
D	25	SER	CYS	engineered mutation	UNP P25779
E	25	SER	CYS	engineered mutation	UNP P25779

- Molecule 2 is N-(1H-benzimidazol-2-yl)-1,3-dimethyl-1H-pyrazole-4-carboxamide (three-letter code: 3H5) (formula: $C_{13}H_{13}N_5O$).
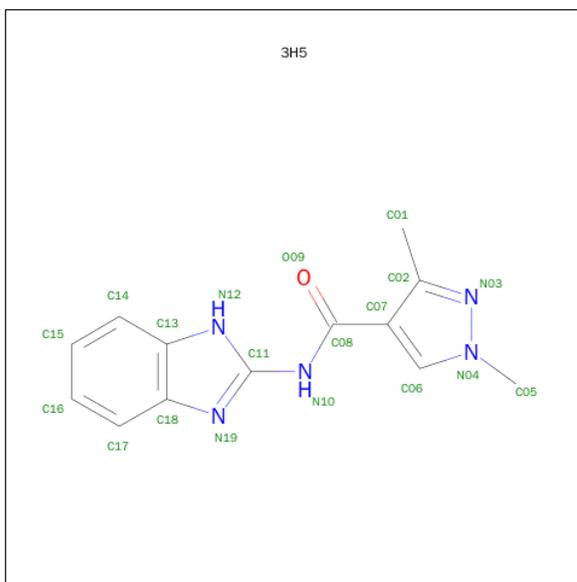
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			19	13	5	1		
2	C	1	Total	C	N	O	0	0
			19	13	5	1		

- Molecule 3 is N-(1H-benzimidazol-2-yl)-3-(4-fluorophenyl)-1H-pyrazole-4-carboxamide (three-letter code: 3H6) (formula: $C_{17}H_{12}FN_5O$).
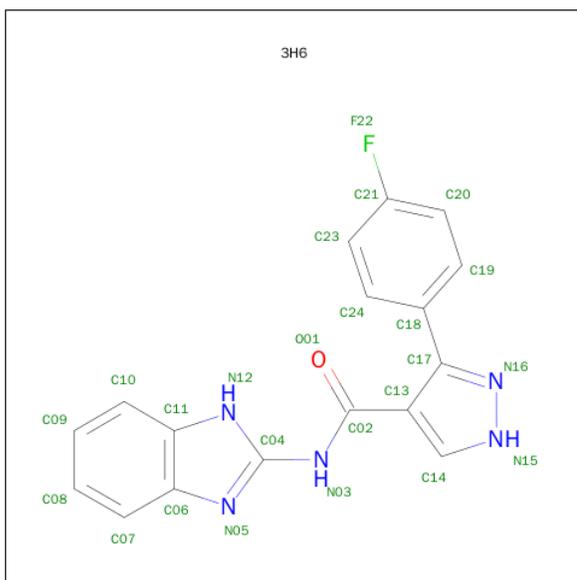
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	F	N	O	0	0
			24	17	1	5	1		

- Molecule 4 is 4,6-difluoro-1,3-benzothiazol-2-amine (three-letter code: 3H7) (formula: $C_7H_4F_2N_2S$).
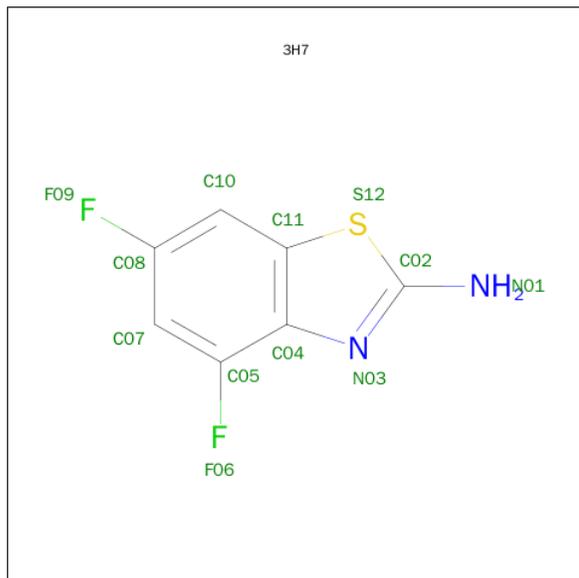
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total 12	C 7	F 2	N 2	S 1	0	0
4	C	1	Total 12	C 7	F 2	N 2	S 1	0	0
4	D	1	Total 12	C 7	F 2	N 2	S 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	6	Total 6	O 6	0	0
5	B	4	Total 4	O 4	0	0
5	C	4	Total 4	O 4	0	0
5	D	4	Total 4	O 4	0	0
5	E	2	Total 2	O 2	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Cruzipain

- Molecule 1: Cruzipain

- Molecule 1: Cruzipain

- Molecule 1: Cruzipain

S143 T148 M152 V156 Q159 L160 G163 L166 Y177 I180 K181 N182 S183 W184 T185 T186 Y193 I194 R195 E207 A212 G215

• **Molecule 1: Cruzipain**

Chain E:

GLY A1 R8 Q21 C22 G23 S24 S25 W26 A27 F28 S29 A30 I31 G32 N33 V34 E35 C36 Q37 W38 H43 L48 S49 E50 Q51 M52 T59 D60 S61 G62 C63 S64 G65 G66 L67 E73 W74 I75 V76 Q77 E78 G81 A82 V83 Y84 S88 Y89 A92 S93

I97 S98 P99 T103 S104 G105 H106 T107 V108 G109 A110 T111 I112 L118 P119 Q120 D121 E122 W128 V135 V139 D140 A141 S142 S143 W144 T148 V156 Q159 H162 G163 V164 L165 L166 V167 W178 I179 I180 Y193 I194 G199 S200 N201 K206 G215

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	137.79Å 137.79Å 166.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.98 – 3.27 29.98 – 3.27	Depositor EDS
% Data completeness (in resolution range)	99.6 (29.98-3.27) 99.7 (29.98-3.27)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.38 (at 3.24Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.8_1069)	Depositor
R, R_{free}	0.193 , 0.236 0.205 , 0.249	Depositor DCC
R_{free} test set	1013 reflections (4.17%)	DCC
Wilson B-factor (Å^2)	84.4	Xtriage
Anisotropy	0.624	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 40.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \lvert L \rvert > = 0.51, < L^2 > = 0.35$	Xtriage
Outliers	0 of 25308 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8074	wwPDB-VP
Average B, all atoms (Å^2)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 3H5, 3H6, 3H7

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	0.47	0/1631	0.62	0/2232
1	B	0.51	0/1631	0.65	0/2232
1	C	0.50	0/1631	0.64	0/2232
1	D	0.45	0/1631	0.64	0/2232
1	E	0.54	0/1627	0.65	0/2228
All	All	0.49	0/8151	0.64	0/11156

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 (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	A	1592	0	1490	17	0
1	B	1592	0	1490	23	0
1	C	1592	0	1490	26	0
1	D	1592	0	1490	28	0
1	E	1588	0	1479	37	0
2	A	19	0	13	1	0
2	C	19	0	13	2	0
3	A	24	0	12	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	12	0	4	2	0
4	C	12	0	4	0	0
4	D	12	0	4	2	0
5	A	6	0	0	0	0
5	B	4	0	0	0	0
5	C	4	0	0	0	0
5	D	4	0	0	0	0
5	E	2	0	0	0	0
All	All	8074	0	7489	130	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 (130) 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)
1:E:25:SER:HB3	1:E:162:HIS:ND1	2.04	0.73
1:B:51:GLN:NE2	1:B:55:SER:OG	2.21	0.72
1:D:25:SER:OG	4:D:301:3H7:N01	2.23	0.71
1:A:185:THR:OG1	1:A:187:GLN:NE2	2.26	0.64
1:B:25:SER:OG	4:B:301:3H7:N01	2.25	0.64
1:A:8:ARG:HD3	1:A:193:TYR:CZ	2.37	0.60
1:E:24:SER:O	1:E:26:TRP:N	2.32	0.59
1:E:74:TRP:HD1	1:E:78:GLU:OE1	1.84	0.59
1:D:166:LEU:HD23	1:D:180:ILE:HG12	1.84	0.59
1:A:8:ARG:HD3	1:A:193:TYR:CE1	2.38	0.59
1:A:71:ALA:O	1:A:75:ILE:HG13	2.03	0.58
1:B:128:TRP:HZ3	1:C:82:ALA:HB2	1.68	0.58
1:E:97:ILE:O	1:E:99:PRO:HD3	2.04	0.58
1:A:159:GLN:HA	1:A:159:GLN:OE1	2.04	0.58
1:C:161:ASP:HB2	2:C:302:3H5:H8	1.87	0.57
1:C:166:LEU:HD23	1:C:180:ILE:HG12	1.85	0.57
1:A:166:LEU:HD23	1:A:180:ILE:HG12	1.87	0.56
1:E:75:ILE:HB	1:E:112:ILE:HD11	1.88	0.56
1:E:75:ILE:HD13	1:E:83:VAL:HG22	1.88	0.56
1:D:36:CYS:SG	1:D:134:PRO:HG2	2.48	0.54
1:C:8:ARG:HD3	1:C:193:TYR:CE1	2.42	0.54
1:E:52:MET:SD	1:E:83:VAL:HA	2.48	0.54
1:D:143:SER:HB3	1:D:156:VAL:HG23	1.90	0.53
1:E:32:GLY:HA2	1:E:35:GLU:OE1	2.09	0.53
2:A:301:3H5:H11	2:A:301:3H5:H4	1.73	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:8:ARG:HD3	1:D:193:TYR:CE1	2.43	0.53
1:E:167:VAL:HG21	1:E:193:TYR:HE1	1.73	0.52
1:E:180:ILE:HB	1:E:194:ILE:HG23	1.91	0.52
1:D:88:SER:O	1:D:106:HIS:ND1	2.39	0.52
1:D:160:LEU:HD11	1:D:207:GLU:HB2	1.92	0.52
1:E:43:HIS:CD2	1:E:43:HIS:N	2.78	0.51
1:A:177:TYR:CE1	1:A:195:ARG:HD2	2.46	0.51
1:D:25:SER:HB3	1:D:163:GLY:H	1.77	0.50
1:D:71:ALA:O	1:D:75:ILE:HG13	2.11	0.50
1:E:8:ARG:HD3	1:E:193:TYR:CE1	2.46	0.50
1:C:180:ILE:HB	1:C:194:ILE:HG23	1.93	0.50
1:A:180:ILE:HB	1:A:194:ILE:HG23	1.94	0.50
1:D:72:PHE:CE1	1:D:212:ALA:HB2	2.47	0.49
1:E:166:LEU:HD23	1:E:180:ILE:HG12	1.94	0.49
1:B:37:GLN:NE2	1:B:132:ASN:O	2.42	0.49
1:C:188:TRP:CH2	1:C:194:ILE:HB	2.47	0.48
1:B:152:MET:HG2	1:B:155:CYS:HB2	1.94	0.48
1:D:25:SER:HG	4:D:301:3H7:H1	1.57	0.48
1:A:138:ALA:HB3	1:A:160:LEU:HD13	1.95	0.48
1:E:33:ASN:O	1:E:37:GLN:HG2	2.12	0.48
1:B:102:THR:OG1	1:B:106:HIS:NE2	2.31	0.48
1:A:36:CYS:SG	1:A:134:PRO:HG2	2.53	0.48
1:D:25:SER:CB	1:D:163:GLY:H	2.27	0.48
1:E:201:ASN:ND2	1:E:206:LYS:HE2	2.30	0.47
1:B:43:HIS:HB3	1:B:44:PRO:HD2	1.95	0.47
1:E:144:TRP:CZ3	1:E:164:VAL:HG11	2.49	0.47
1:B:53:LEU:O	1:B:57:ASP:HB3	2.16	0.46
1:C:26:TRP:CG	1:C:66:GLY:HA3	2.50	0.46
1:C:97:ILE:O	1:C:99:PRO:HD3	2.15	0.46
1:E:26:TRP:CE3	1:E:27:ALA:HB2	2.52	0.45
1:E:38:TRP:CZ2	1:E:43:HIS:ND1	2.85	0.45
1:B:114:GLY:HA3	1:C:103:THR:HB	1.98	0.45
1:C:48:LEU:O	1:C:86:GLU:HB2	2.16	0.45
1:B:180:ILE:HB	1:B:194:ILE:HG23	1.99	0.45
1:D:39:PHE:HA	1:D:43:HIS:O	2.17	0.45
1:D:75:ILE:HD13	1:D:83:VAL:HG22	1.97	0.45
1:C:118:LEU:HB2	1:C:209:ALA:HB3	1.99	0.45
1:E:88:SER:O	1:E:106:HIS:ND1	2.49	0.45
1:C:12:ALA:HA	1:C:39:PHE:CD1	2.52	0.45
1:E:74:TRP:CD1	1:E:78:GLU:OE1	2.69	0.44
1:E:92:ALA:HB3	1:E:99:PRO:HG2	1.99	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:31:ILE:O	1:D:35:GLU:HG3	2.18	0.44
1:D:8:ARG:HD3	1:D:193:TYR:CZ	2.52	0.44
1:C:31:ILE:O	1:C:35:GLU:HG3	2.17	0.44
1:D:31:ILE:HA	1:D:31:ILE:HD13	1.85	0.44
1:B:57:ASP:HA	1:B:74:TRP:CE2	2.53	0.44
1:E:118:LEU:HD11	1:E:128:TRP:CD1	2.53	0.44
1:B:93:SER:HA	1:B:97:ILE:O	2.18	0.44
1:A:97:ILE:O	1:A:99:PRO:HD3	2.18	0.43
1:B:51:GLN:HB2	1:B:91:TYR:HA	2.00	0.43
1:E:84:TYR:HA	1:E:107:THR:O	2.17	0.43
1:B:31:ILE:HG13	1:B:50:GLU:HA	2.01	0.43
1:D:12:ALA:HA	1:D:39:PHE:CD1	2.53	0.43
1:E:140:ASP:OD2	1:E:156:VAL:HG12	2.18	0.43
1:D:97:ILE:O	1:D:99:PRO:HD3	2.19	0.43
3:A:302:3H6:H4	3:A:302:3H6:C19	2.31	0.43
1:E:24:SER:C	1:E:26:TRP:N	2.72	0.43
1:A:115:HIS:O	1:B:105:GLY:HA2	2.18	0.43
1:E:25:SER:O	1:E:163:GLY:N	2.51	0.43
1:A:25:SER:HB3	1:A:162:HIS:CE1	2.54	0.43
1:E:120:GLN:HA	1:E:206:LYS:O	2.19	0.42
1:C:31:ILE:HG23	1:C:48:LEU:HB2	2.01	0.42
1:E:122:GLU:HG2	1:E:178:TRP:HE1	1.84	0.42
1:C:71:ALA:O	1:C:75:ILE:HG13	2.19	0.42
1:E:28:PHE:CG	1:E:50:GLU:HG2	2.54	0.42
1:A:182:ASN:OD1	1:A:183:SER:N	2.53	0.42
1:D:51:GLN:NE2	1:D:55:SER:OG	2.45	0.42
1:B:16:VAL:HG21	1:B:192:GLY:HA3	2.01	0.42
1:D:28:PHE:CG	1:D:50:GLU:HG2	2.55	0.42
1:B:118:LEU:HD23	1:B:118:LEU:HA	1.84	0.42
1:D:180:ILE:HB	1:D:194:ILE:HG23	2.01	0.42
1:E:74:TRP:CD1	1:E:78:GLU:HB2	2.54	0.42
1:E:139:VAL:HG21	1:E:144:TRP:CH2	2.55	0.42
1:B:75:ILE:HD13	1:B:83:VAL:HG22	2.02	0.42
1:B:31:ILE:HG23	1:B:48:LEU:HB2	2.01	0.41
1:B:160:LEU:HD23	1:B:160:LEU:HA	1.81	0.41
1:D:17:LYS:NZ	1:D:35:GLU:OE2	2.35	0.41
1:C:166:LEU:CD2	1:C:180:ILE:HG12	2.49	0.41
2:C:302:3H5:H11	2:C:302:3H5:H4	1.84	0.41
1:A:128:TRP:HZ3	1:B:82:ALA:HB2	1.86	0.41
1:D:177:TYR:CE1	1:D:195:ARG:HD2	2.55	0.41
1:D:92:ALA:C	1:D:94:GLY:H	2.24	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:63:CYS:SG	1:D:93:SER:HB3	2.60	0.41
1:C:168:GLY:O	1:C:179:ILE:N	2.38	0.41
1:B:147:TYR:CE1	1:B:194:ILE:HG13	2.56	0.41
1:D:182:ASN:OD1	1:D:183:SER:N	2.52	0.41
1:A:138:ALA:CB	1:A:160:LEU:HD13	2.51	0.41
1:E:135:VAL:O	1:E:165:LEU:HD12	2.20	0.41
1:C:31:ILE:HG13	1:C:50:GLU:HA	2.03	0.41
1:E:122:GLU:CD	1:E:199:GLY:H	2.24	0.41
1:C:89:TYR:CD1	1:C:106:HIS:CE1	3.08	0.41
1:C:36:CYS:SG	1:C:134:PRO:HG2	2.61	0.41
1:A:1:ALA:HB1	1:A:169:TYR:OH	2.20	0.41
1:B:25:SER:CB	4:B:301:3H7:H1	2.29	0.41
1:D:159:GLN:HA	1:D:159:GLN:OE1	2.20	0.40
1:C:24:SER:HA	1:C:26:TRP:CD1	2.56	0.40
1:C:31:ILE:HA	1:C:31:ILE:HD13	1.85	0.40
1:C:63:CYS:SG	1:C:93:SER:HB3	2.61	0.40
1:C:182:ASN:OD1	1:C:183:SER:N	2.52	0.40
1:E:62:GLY:O	1:E:63:CYS:HB2	2.22	0.40
1:C:61:SER:HA	1:C:62:GLY:HA3	1.84	0.40
1:E:81:GLY:O	1:E:111:THR:HA	2.22	0.40
1:C:147:TYR:CZ	1:C:194:ILE:HG13	2.57	0.40
1:E:52:MET:HE2	1:E:83:VAL:HG13	2.03	0.40
1:E:31:ILE:HG23	1:E:48:LEU:HB2	2.04	0.40

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 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	213/216 (99%)	194 (91%)	16 (8%)	3 (1%)	14	54
1	B	213/216 (99%)	195 (92%)	18 (8%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	213/216 (99%)	194 (91%)	18 (8%)	1 (0%)	34	74
1	D	213/216 (99%)	191 (90%)	19 (9%)	3 (1%)	14	54
1	E	213/216 (99%)	190 (89%)	20 (9%)	3 (1%)	14	54
All	All	1065/1080 (99%)	964 (90%)	91 (8%)	10 (1%)	21	63

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	25	SER
1	A	62	GLY
1	D	62	GLY
1	E	26	TRP
1	A	93	SER
1	E	93	SER
1	A	99	PRO
1	D	185	THR
1	C	62	GLY
1	D	99	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	A	167/167 (100%)	152 (91%)	15 (9%)	12	42
1	B	167/167 (100%)	155 (93%)	12 (7%)	18	55
1	C	167/167 (100%)	153 (92%)	14 (8%)	14	47
1	D	167/167 (100%)	156 (93%)	11 (7%)	21	60
1	E	166/167 (99%)	156 (94%)	10 (6%)	24	64
All	All	834/835 (100%)	772 (93%)	62 (7%)	17	54

All (62) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	SER
1	A	61	SER
1	A	63	CYS
1	A	64	SER
1	A	67	LEU
1	A	73	GLU
1	A	80	ASN
1	A	89	TYR
1	A	101	CYS
1	A	102	THR
1	A	103	THR
1	A	120	GLN
1	A	143	SER
1	A	159	GLN
1	A	187	GLN
1	B	25	SER
1	B	61	SER
1	B	67	LEU
1	B	73	GLU
1	B	80	ASN
1	B	87	ASP
1	B	89	TYR
1	B	115	HIS
1	B	148	THR
1	B	159	GLN
1	B	185	THR
1	B	191	GLU
1	C	25	SER
1	C	61	SER
1	C	63	CYS
1	C	64	SER
1	C	67	LEU
1	C	73	GLU
1	C	80	ASN
1	C	89	TYR
1	C	98	SER
1	C	104	SER
1	C	120	GLN
1	C	143	SER
1	C	159	GLN
1	C	214	VAL
1	D	61	SER
1	D	63	CYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	64	SER
1	D	67	LEU
1	D	88	SER
1	D	89	TYR
1	D	97	ILE
1	D	120	GLN
1	D	122	GLU
1	D	148	THR
1	D	186	THR
1	E	25	SER
1	E	61	SER
1	E	63	CYS
1	E	64	SER
1	E	67	LEU
1	E	89	TYR
1	E	97	ILE
1	E	120	GLN
1	E	148	THR
1	E	159	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	187	GLN
1	E	202	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

6 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	#$|Z| > 2$
2	3H5	A	301	-	19,21,21	2.02	3 (15%)	14,30,30	2.13	4 (28%)
3	3H6	A	302	-	26,27,27	2.90	7 (26%)	24,38,38	2.09	4 (16%)
4	3H7	B	301	-	10,13,13	1.71	2 (20%)	7,19,19	3.07	5 (71%)
4	3H7	C	301	-	10,13,13	1.87	3 (30%)	7,19,19	2.75	4 (57%)
2	3H5	C	302	-	19,21,21	1.97	2 (10%)	14,30,30	2.67	5 (35%)
4	3H7	D	301	-	10,13,13	1.64	3 (30%)	7,19,19	2.41	3 (42%)

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
2	3H5	A	301	-	-	0/4/8/8	0/3/3/3
3	3H6	A	302	-	-	0/8/12/12	0/4/4/4
4	3H7	B	301	-	-	0/0/0/0	0/2/2/2
4	3H7	C	301	-	-	0/0/0/0	0/2/2/2
2	3H5	C	302	-	-	0/4/8/8	0/3/3/3
4	3H7	D	301	-	-	0/0/0/0	0/2/2/2

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	302	3H6	C18-C17	-12.64	1.35	1.49
4	D	301	3H7	C10-C11	2.01	1.40	1.37
3	A	302	3H6	C08-C07	2.10	1.41	1.36
3	A	302	3H6	C14-N15	2.11	1.38	1.33
3	A	302	3H6	C08-C09	2.22	1.43	1.38
2	A	301	3H5	C15-C14	2.23	1.41	1.36
4	B	301	3H7	C10-C08	2.26	1.39	1.35

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	302	3H6	C10-C11	2.45	1.46	1.41
4	D	301	3H7	C10-C08	2.48	1.40	1.35
4	C	301	3H7	C10-C11	2.61	1.41	1.37
4	D	301	3H7	C02-N01	2.79	1.43	1.35
4	C	301	3H7	C10-C08	2.86	1.40	1.35
4	C	301	3H7	C02-N01	2.94	1.43	1.35
4	B	301	3H7	C02-N01	2.96	1.43	1.35
3	A	302	3H6	C23-C24	3.06	1.44	1.38
3	A	302	3H6	C02-N03	3.09	1.43	1.35
2	A	301	3H5	C08-N10	3.81	1.45	1.35
2	C	302	3H5	C08-N10	3.88	1.45	1.35
2	C	302	3H5	C11-N10	6.01	1.45	1.39
2	A	301	3H5	C11-N10	6.43	1.46	1.39

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	302	3H5	C11-N10-C08	-7.24	114.58	128.18
3	A	302	3H6	C04-N03-C02	-5.81	117.26	128.18
2	A	301	3H5	C11-N10-C08	-4.30	120.10	128.18
4	B	301	3H7	C07-C05-C04	-3.43	119.20	123.88
4	C	301	3H7	C07-C05-C04	-2.85	120.00	123.88
4	D	301	3H7	C07-C05-C04	-2.69	120.22	123.88
3	A	302	3H6	O01-C02-N03	-2.49	118.22	123.68
2	A	301	3H5	O09-C08-N10	-2.43	118.34	123.68
2	C	302	3H5	O09-C08-N10	-2.38	118.46	123.68
2	C	302	3H5	C01-C02-N03	2.22	124.75	119.87
4	C	301	3H7	F09-C08-C07	2.24	121.20	118.22
4	B	301	3H7	F06-C05-C04	2.52	123.13	118.18
2	A	301	3H5	C07-C08-N10	2.64	121.48	116.13
4	B	301	3H7	F09-C08-C07	2.68	121.79	118.22
4	D	301	3H7	N01-C02-N03	3.56	127.58	122.92
3	A	302	3H6	C13-C02-N03	3.59	123.41	116.13
2	C	302	3H5	C07-C08-N10	3.64	123.51	116.13
4	D	301	3H7	C05-C07-C08	3.79	120.59	116.62
2	C	302	3H5	C02-N03-N04	4.01	107.92	104.52
4	C	301	3H7	N01-C02-N03	4.17	128.38	122.92
4	C	301	3H7	C05-C07-C08	4.28	121.10	116.62
4	B	301	3H7	C05-C07-C08	4.33	121.16	116.62
4	B	301	3H7	N01-C02-N03	4.66	129.02	122.92
2	A	301	3H5	C02-N03-N04	4.90	108.68	104.52
3	A	302	3H6	C18-C17-N16	5.52	127.83	120.71

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	3H5	1	0
3	A	302	3H6	1	0
4	B	301	3H7	2	0
2	C	302	3H5	2	0
4	D	301	3H7	2	0

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)

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	215/216 (99%)	0.04	10 (4%)	35	28	62, 82, 115, 141	0
1	B	215/216 (99%)	-0.07	4 (1%)	70	61	59, 80, 115, 142	0
1	C	215/216 (99%)	-0.11	4 (1%)	70	61	60, 84, 117, 142	0
1	D	215/216 (99%)	-0.07	3 (1%)	78	71	62, 85, 117, 142	0
1	E	215/216 (99%)	0.40	17 (7%)	15	11	72, 88, 122, 154	0
All	All	1075/1080 (99%)	0.04	38 (3%)	48	39	59, 85, 118, 154	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	172	SER	4.0
1	E	104	SER	3.8
1	E	97	ILE	3.7
1	D	215	GLY	3.7
1	B	97	ILE	3.6
1	E	215	GLY	3.5
1	E	142	SER	3.3
1	E	143	SER	3.0
1	E	59	THR	3.0
1	E	103	THR	3.0
1	C	97	ILE	3.0
1	E	23	GLY	2.8
1	A	33	ASN	2.6
1	E	29	SER	2.6
1	B	29	SER	2.6
1	A	25	SER	2.5
1	A	29	SER	2.5
1	A	3	ALA	2.5
1	D	97	ILE	2.5
1	E	111	THR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	65	GLY	2.4
1	E	108	VAL	2.4
1	E	109	GLY	2.4
1	A	175	VAL	2.2
1	E	73	GLU	2.2
1	D	152	MET	2.2
1	A	174	ALA	2.2
1	C	174	ALA	2.2
1	E	21	GLN	2.2
1	B	163	GLY	2.2
1	C	29	SER	2.1
1	A	173	ALA	2.1
1	E	92	ALA	2.1
1	A	97	ILE	2.1
1	B	95	GLU	2.1
1	E	76	VAL	2.1
1	A	32	GLY	2.0
1	C	100	PRO	2.0

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

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

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
4	3H7	C	301	12/12	0.67	0.46	4.46	62,114,132,152	12
4	3H7	D	301	12/12	0.67	0.45	3.73	90,108,116,138	12
2	3H5	A	301	19/19	0.76	0.47	2.74	77,104,121,132	19

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å²)	Q<0.9
3	3H6	A	302	24/24	0.73	0.46	2.49	71,109,147,161	24
2	3H5	C	302	19/19	0.79	0.34	2.01	80,108,118,124	19
4	3H7	B	301	12/12	0.81	0.21	-0.36	58,109,133,134	1

6.5 Other polymers (i)

There are no such residues in this entry.