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

Full wwPDB X-ray Structure Validation Report (i)

Aug 21, 2020 – 03:09 AM BST

PDB ID	:	4TYA
Title	:	An Ligand-observed Mass Spectrometry-based Approach Integrated into the Fragment Based Lead Discovery Pipeline
Authors	:	Shui, W.; Yang, C.; Lin, J.; Chen, X.; Qin, S.; Chen, S.
Deposited on	:	2014-07-08
Resolution	:	2.94 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 2.94 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		0.267
Clashscore		5
Ramachandran outliers		0.2%
Sidechain outliers		6.5%
RSRZ outliers		0.5%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
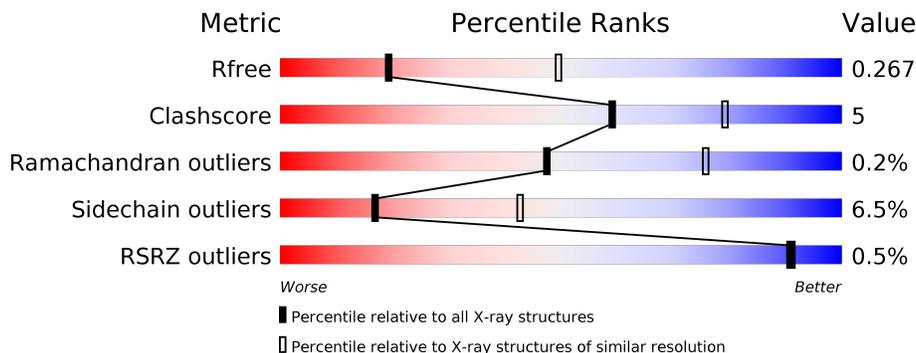
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2969 (2.98-2.90)
Clashscore	141614	3218 (2.98-2.90)
Ramachandran outliers	138981	3122 (2.98-2.90)
Sidechain outliers	138945	3124 (2.98-2.90)
RSRZ outliers	127900	2902 (2.98-2.90)

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% 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	566	80% 15% · ·
1	B	566	% 79% 15% · ·
1	C	566	82% 13% · ·
1	D	566	83% 11% · ·

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 16940 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 Polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	544	Total	C	N	O	S	0	0	0
			4222	2664	744	783	31			
1	B	544	Total	C	N	O	S	0	0	0
			4222	2664	744	783	31			
1	C	544	Total	C	N	O	S	0	0	0
			4222	2664	744	783	31			
1	D	544	Total	C	N	O	S	0	0	0
			4222	2664	744	783	31			

- Molecule 2 is 4-(trifluoromethyl)benzoic acid (three-letter code: 3AE) (formula: $C_8H_5F_3O_2$).
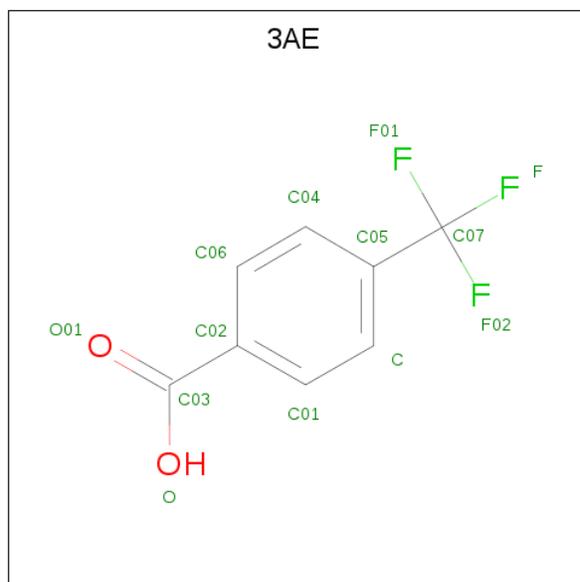
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	F	O	0	0
			13	8	3	2		
2	B	1	Total	C	F	O	0	0
			13	8	3	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	C	1	Total 13	C 8	F 3	O 2	0	0
2	D	1	Total 13	C 8	F 3	O 2	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry 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: Polyprotein

Chain A:
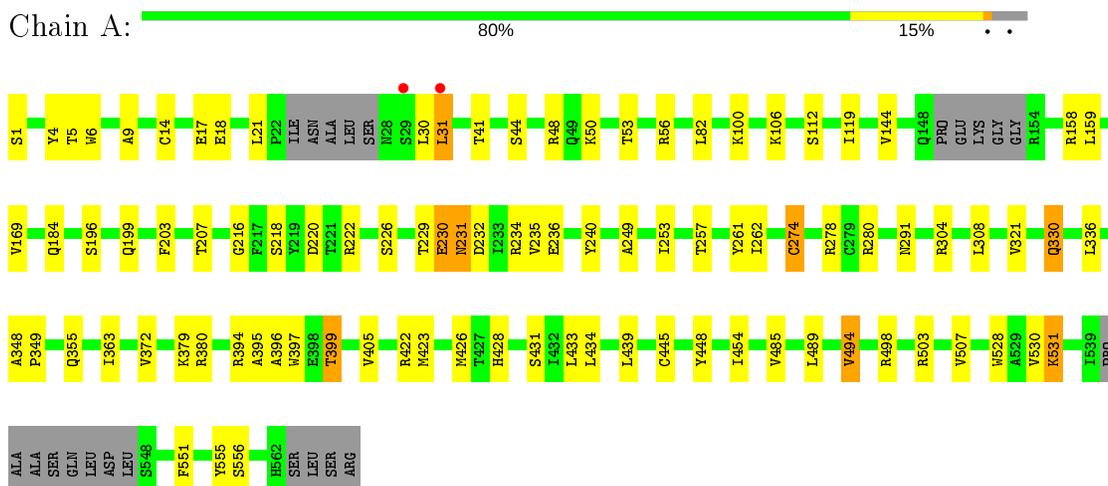
• Molecule 1: Polyprotein
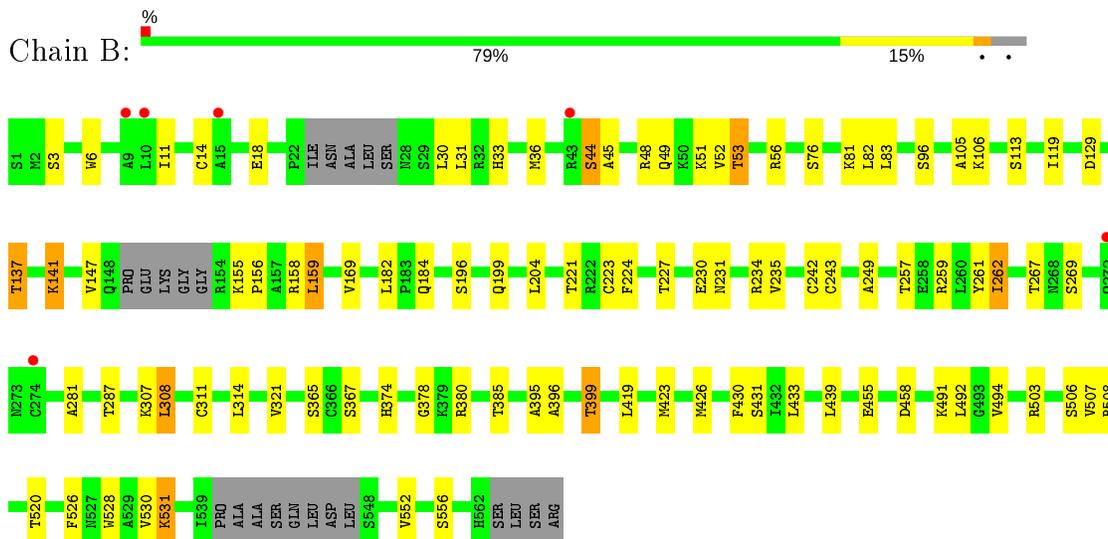
Chain B:

• Molecule 1: Polyprotein
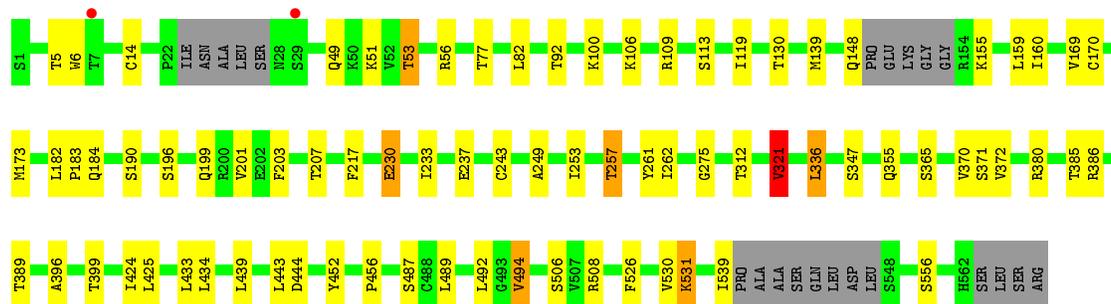
Chain C:

- Molecule 1: Polyprotein

Chain D:
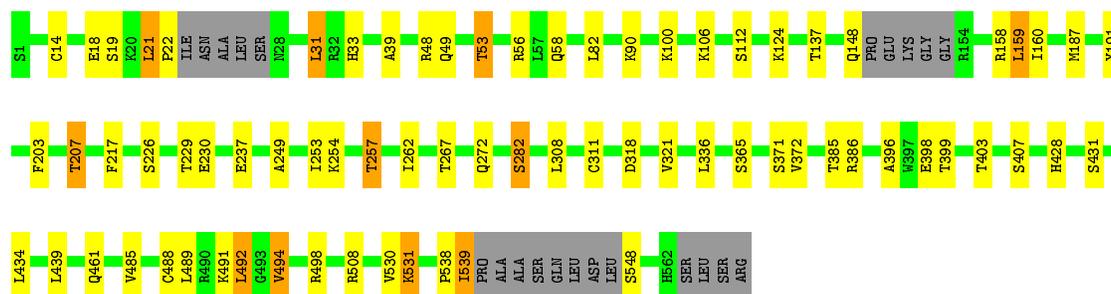
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	102.18Å 102.16Å 251.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.34 – 2.94 44.96 – 2.94	Depositor EDS
% Data completeness (in resolution range)	98.4 (47.34-2.94) 98.4 (44.96-2.94)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.71 (at 2.96Å)	Xtriage
Refinement program	REFMAC 5.8.0049, REFMAC 5.8.0049	Depositor
R, R_{free}	0.201 , 0.270 0.205 , 0.267	Depositor DCC
R_{free} test set	2832 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	38.4	Xtriage
Anisotropy	0.067	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 17.4	EDS
L-test for twinning [2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.087 for k,h,-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	16940	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.57% 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: 3AE

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	0.54	0/4313	0.74	1/5852 (0.0%)
1	B	0.53	0/4313	0.76	0/5852
1	C	0.53	0/4313	0.77	1/5852 (0.0%)
1	D	0.53	0/4313	0.74	2/5852 (0.0%)
All	All	0.53	0/17252	0.75	4/23408 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	21	LEU	CA-CB-CG	5.25	127.37	115.30
1	D	31	LEU	CA-CB-CG	5.21	127.29	115.30
1	C	321	VAL	CB-CA-C	-5.08	101.75	111.40
1	A	304	ARG	NE-CZ-NH1	5.07	122.84	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	548	SER	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	A	4222	0	4212	41	0
1	B	4222	0	4212	46	0
1	C	4222	0	4212	37	0
1	D	4222	0	4212	35	0
2	A	13	0	4	0	0
2	B	13	0	4	0	0
2	C	13	0	4	0	0
2	D	13	0	4	0	0
All	All	16940	0	16864	158	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 5.

All (158) 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:D:49:GLN:O	1:D:53:THR:HG22	1.63	0.98
1:C:49:GLN:O	1:C:53:THR:HG22	1.64	0.97
1:B:49:GLN:O	1:B:53:THR:HG23	1.72	0.88
1:C:396:ALA:O	1:C:399:THR:HG22	1.71	0.88
1:A:395:ALA:O	1:A:399:THR:HG22	1.80	0.81
1:B:49:GLN:O	1:B:53:THR:CG2	2.29	0.81
1:B:395:ALA:O	1:B:399:THR:HG22	1.85	0.77
1:A:6:TRP:HE1	1:A:53:THR:HG21	1.48	0.77
1:D:489:LEU:HA	1:D:494:VAL:HG13	1.68	0.74
1:D:203:PHE:O	1:D:207:THR:HG23	1.88	0.73
1:C:82:LEU:HD13	1:C:249:ALA:HB2	1.70	0.73
1:A:253:ILE:O	1:A:257:THR:HG23	1.88	0.73
1:B:433:LEU:HB3	1:B:439:LEU:HD23	1.71	0.73
1:D:253:ILE:O	1:D:257:THR:HG23	1.89	0.72
1:C:508:ARG:HG3	1:C:526:PHE:HB2	1.72	0.72
1:C:196:SER:H	1:C:199:GLN:HE21	1.43	0.67
1:B:257:THR:HG22	1:B:261:TYR:HD2	1.60	0.66
1:D:217:PHE:CD1	1:D:336:LEU:HD21	2.32	0.65
1:C:92:THR:O	1:C:109:ARG:NH1	2.30	0.65

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:445:CYS:SG	1:A:454:ILE:HD12	2.38	0.64
1:A:203:PHE:O	1:A:207:THR:HG23	1.98	0.63
1:A:82:LEU:HD13	1:A:249:ALA:HB2	1.81	0.63
1:A:31:LEU:HD23	1:A:396:ALA:HB2	1.81	0.63
1:B:30:LEU:O	1:B:494:VAL:HB	1.98	0.62
1:B:196:SER:H	1:B:199:GLN:HE21	1.48	0.62
1:D:321:VAL:HG21	1:D:365:SER:HB3	1.82	0.62
1:D:49:GLN:O	1:D:53:THR:CG2	2.43	0.62
1:B:48:ARG:HG3	1:B:159:LEU:HD22	1.80	0.61
1:A:48:ARG:HG2	1:A:159:LEU:HD13	1.82	0.60
1:C:82:LEU:CD1	1:C:249:ALA:HB2	2.31	0.60
1:B:234:ARG:HG2	1:B:262:ILE:HD11	1.82	0.60
1:D:396:ALA:O	1:D:399:THR:HG22	2.00	0.60
1:D:539:ILE:HD13	1:D:539:ILE:N	2.19	0.57
1:A:1:SER:OG	1:A:230:GLU:OE1	2.17	0.56
1:B:530:VAL:O	1:B:531:LYS:CB	2.53	0.56
1:C:530:VAL:O	1:C:531:LYS:CB	2.53	0.56
1:B:508:ARG:HG3	1:B:526:PHE:HB2	1.88	0.56
1:B:281:ALA:O	1:B:287:THR:OG1	2.24	0.56
1:A:433:LEU:HB3	1:A:439:LEU:HD23	1.87	0.55
1:D:434:LEU:CD1	1:D:439:LEU:HD11	2.36	0.55
1:B:33:HIS:HD2	1:B:491:LYS:O	1.89	0.55
1:D:321:VAL:CG2	1:D:365:SER:HB3	2.37	0.54
1:C:203:PHE:O	1:C:207:THR:HG23	2.07	0.54
1:C:253:ILE:O	1:C:257:THR:HG23	2.08	0.54
1:B:129:ASP:O	1:B:259:ARG:HD2	2.07	0.54
1:D:48:ARG:HG2	1:D:159:LEU:HD13	1.89	0.54
1:A:4:TYR:HB2	1:A:53:THR:HG22	1.91	0.53
1:D:187:MET:HE3	1:D:191:TYR:CD1	2.44	0.52
1:D:398:GLU:OE2	1:D:407:SER:OG	2.18	0.52
1:D:18:GLU:HG2	1:D:22:PRO:HG3	1.91	0.52
1:C:385:THR:OG1	1:C:386:ARG:N	2.43	0.52
1:A:30:LEU:O	1:A:494:VAL:HB	2.11	0.51
1:D:321:VAL:HG21	1:D:365:SER:CB	2.40	0.51
1:B:31:LEU:HD22	1:B:492:LEU:HD23	1.92	0.50
1:D:237:GLU:CG	1:D:257:THR:HG21	2.42	0.50
1:C:489:LEU:HA	1:C:494:VAL:CG1	2.41	0.50
1:B:396:ALA:O	1:B:399:THR:HG23	2.11	0.50
1:C:217:PHE:CD1	1:C:336:LEU:HD21	2.46	0.50
1:B:6:TRP:HE1	1:B:53:THR:HG21	1.77	0.49
1:C:6:TRP:HE1	1:C:53:THR:HG21	1.77	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:428:HIS:O	1:A:431:SER:HB3	2.11	0.49
1:D:488:CYS:O	1:D:492:LEU:HB2	2.13	0.49
1:A:5:THR:HG23	1:A:278:ARG:NH2	2.28	0.49
1:A:330:GLN:HA	1:A:330:GLN:HE21	1.78	0.48
1:A:489:LEU:HA	1:A:494:VAL:HG13	1.95	0.48
1:B:503:ARG:O	1:B:507:VAL:HG23	2.13	0.48
1:B:82:LEU:HD13	1:B:249:ALA:HB2	1.95	0.48
1:D:308:LEU:HB2	1:D:311:CYS:SG	2.53	0.48
1:D:371:SER:OG	1:D:372:VAL:N	2.47	0.48
1:D:385:THR:OG1	1:D:386:ARG:N	2.47	0.48
1:C:233:ILE:HB	1:C:262:ILE:HD12	1.95	0.48
1:A:434:LEU:HD13	1:A:439:LEU:HD11	1.96	0.47
1:C:207:THR:HG21	1:C:312:THR:HG21	1.97	0.47
1:B:426:MET:O	1:B:430:PHE:HB2	2.14	0.47
1:B:204:LEU:HD21	1:B:314:LEU:HD23	1.95	0.47
1:B:231:ASN:O	1:B:235:VAL:HG23	2.15	0.47
1:C:443:LEU:HD12	1:C:456:PRO:HG3	1.97	0.47
1:A:485:VAL:O	1:A:489:LEU:HG	2.15	0.47
1:A:4:TYR:HB2	1:A:53:THR:CG2	2.45	0.47
1:A:280:ARG:NE	1:A:291:ASN:OD1	2.42	0.47
1:A:530:VAL:O	1:A:531:LYS:CB	2.63	0.46
1:C:257:THR:HA	1:C:261:TYR:HB2	1.97	0.46
1:C:49:GLN:O	1:C:53:THR:CG2	2.51	0.46
1:D:58:GLN:NE2	1:D:229:THR:OG1	2.48	0.46
1:B:182:LEU:HD13	1:B:243:CYS:SG	2.55	0.46
1:A:503:ARG:O	1:A:507:VAL:HG23	2.15	0.46
1:D:19:SER:HA	1:D:39:ALA:HB3	1.98	0.46
1:D:33:HIS:HD2	1:D:491:LYS:O	1.99	0.46
1:B:119:ILE:HD13	1:B:169:VAL:HG11	1.97	0.46
1:B:96:SER:HB3	1:B:105:ALA:HB2	1.98	0.46
1:B:76:SER:HA	1:B:242:CYS:O	2.15	0.46
1:A:17:GLU:HB3	1:A:41:THR:HG22	1.98	0.45
1:A:348:ALA:N	1:A:349:PRO:CD	2.79	0.45
1:A:6:TRP:NE1	1:A:53:THR:HG21	2.25	0.45
1:A:9:ALA:HB3	1:A:274:CYS:HA	1.98	0.45
1:C:530:VAL:O	1:C:531:LYS:HB2	2.16	0.45
1:A:236:GLU:HG2	1:A:240:TYR:CZ	2.51	0.45
1:A:448:TYR:CE2	1:A:551:PHE:HD1	2.34	0.45
1:A:423:MET:HA	1:A:528:TRP:CZ2	2.52	0.45
1:B:307:LYS:NZ	1:C:148:GLN:OE1	2.48	0.45
1:B:321:VAL:CG2	1:B:365:SER:HB3	2.47	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:3:SER:OG	1:B:53:THR:HA	2.17	0.45
1:D:508:ARG:NE	1:D:530:VAL:HG22	2.31	0.45
1:A:196:SER:OG	1:A:199:GLN:HG3	2.17	0.44
1:C:230:GLU:O	1:C:262:ILE:HD11	2.16	0.44
1:D:321:VAL:CG2	1:D:365:SER:CB	2.95	0.44
1:A:231:ASN:O	1:A:235:VAL:HG23	2.17	0.44
1:C:433:LEU:HB3	1:C:439:LEU:HD23	2.00	0.44
1:D:217:PHE:CE1	1:D:336:LEU:HD21	2.52	0.44
1:A:21:LEU:HD13	1:A:397:TRP:HA	2.00	0.43
1:D:160:ILE:HG13	1:D:282:SER:HB2	2.00	0.43
1:B:455:GLU:HB2	1:B:458:ASP:OD2	2.18	0.43
1:C:389:THR:HG23	1:C:492:LEU:HD11	1.99	0.43
1:C:237:GLU:HG3	1:C:257:THR:HG21	2.01	0.43
1:C:489:LEU:HA	1:C:494:VAL:HG13	1.99	0.43
1:B:423:MET:HA	1:B:528:TRP:CZ2	2.53	0.43
1:B:321:VAL:HG22	1:B:365:SER:HB3	2.01	0.43
1:A:422:ARG:O	1:A:426:MET:HB2	2.18	0.43
1:B:257:THR:HG22	1:B:261:TYR:CD2	2.48	0.43
1:D:530:VAL:O	1:D:531:LYS:CB	2.67	0.42
1:A:144:VAL:HB	1:A:394:ARG:HG2	2.00	0.42
1:B:321:VAL:CG2	1:B:365:SER:CB	2.97	0.42
1:C:170:CYS:HA	1:C:173:MET:CE	2.50	0.42
1:A:257:THR:HA	1:A:261:TYR:HB2	2.01	0.42
1:B:44:SER:HB2	1:B:156:PRO:HA	2.02	0.42
1:A:196:SER:H	1:A:199:GLN:HE21	1.67	0.42
1:B:137:THR:HA	1:B:267:THR:O	2.19	0.42
1:B:257:THR:O	1:B:262:ILE:HB	2.19	0.42
1:B:11:ILE:HD12	1:B:45:ALA:HB1	2.01	0.42
1:C:424:ILE:O	1:C:425:LEU:C	2.57	0.42
1:D:187:MET:CE	1:D:191:TYR:CD1	3.03	0.42
1:C:182:LEU:HD13	1:C:243:CYS:SG	2.60	0.41
1:D:538:PRO:C	1:D:539:ILE:HD13	2.40	0.41
1:A:119:ILE:HD13	1:A:169:VAL:HG11	2.02	0.41
1:B:221:THR:HG21	1:B:224:PHE:HD2	1.85	0.41
1:B:308:LEU:HB3	1:B:311:CYS:SG	2.61	0.41
1:D:428:HIS:O	1:D:431:SER:HB3	2.21	0.41
1:A:234:ARG:HG3	1:A:262:ILE:HD11	2.02	0.41
1:D:82:LEU:HD13	1:D:249:ALA:HB2	2.03	0.41
1:A:229:THR:O	1:A:232:ASP:HB2	2.21	0.41
1:C:170:CYS:HA	1:C:173:MET:HE3	2.02	0.41
1:C:530:VAL:HG12	1:C:531:LYS:N	2.35	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:488:CYS:SG	1:D:492:LEU:HD22	2.60	0.41
1:A:216:GLY:HA3	1:A:363:ILE:HD11	2.01	0.41
1:A:555:TYR:O	1:A:556:SER:C	2.59	0.41
1:C:182:LEU:N	1:C:183:PRO:CD	2.84	0.41
1:C:5:THR:O	1:C:275:GLY:HA3	2.21	0.41
1:C:321:VAL:HG22	1:C:365:SER:CB	2.51	0.41
1:B:30:LEU:HD23	1:B:396:ALA:HA	2.03	0.41
1:B:48:ARG:O	1:B:52:VAL:HG22	2.20	0.41
1:C:119:ILE:HD13	1:C:169:VAL:HG11	2.02	0.41
1:C:201:VAL:HG22	1:C:370:VAL:HG22	2.02	0.41
1:D:485:VAL:O	1:D:489:LEU:HG	2.21	0.41
1:B:33:HIS:HB3	1:B:36:MET:HG3	2.03	0.40
1:B:374:HIS:HD2	1:B:378:GLY:O	2.03	0.40
1:B:141:LYS:HE2	1:B:158:ARG:NH1	2.36	0.40
1:B:223:CYS:O	1:B:227:THR:HG23	2.22	0.40
1:C:444:ASP:HA	1:C:452:TYR:O	2.22	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	536/566 (95%)	511 (95%)	24 (4%)	1 (0%)	47	76
1	B	536/566 (95%)	506 (94%)	29 (5%)	1 (0%)	47	76
1	C	536/566 (95%)	504 (94%)	30 (6%)	2 (0%)	34	64
1	D	536/566 (95%)	506 (94%)	29 (5%)	1 (0%)	47	76
All	All	2144/2264 (95%)	2027 (94%)	112 (5%)	5 (0%)	47	76

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	531	LYS
1	B	531	LYS
1	C	531	LYS
1	D	531	LYS
1	C	347	SER

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	460/482 (95%)	430 (94%)	30 (6%)	17	43
1	B	460/482 (95%)	430 (94%)	30 (6%)	17	43
1	C	460/482 (95%)	431 (94%)	29 (6%)	18	44
1	D	460/482 (95%)	430 (94%)	30 (6%)	17	43
All	All	1840/1928 (95%)	1721 (94%)	119 (6%)	17	43

All (119) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	CYS
1	A	18	GLU
1	A	31	LEU
1	A	44	SER
1	A	50	LYS
1	A	56	ARG
1	A	100	LYS
1	A	106	LYS
1	A	112	SER
1	A	158	ARG
1	A	184	GLN
1	A	218	SER
1	A	220	ASP
1	A	222	ARG
1	A	226	SER
1	A	230	GLU
1	A	231	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	274	CYS
1	A	308	LEU
1	A	321	VAL
1	A	330	GLN
1	A	336	LEU
1	A	355	GLN
1	A	372	VAL
1	A	379	LYS
1	A	380	ARG
1	A	399	THR
1	A	405	VAL
1	A	494	VAL
1	A	498	ARG
1	B	14	CYS
1	B	18	GLU
1	B	44	SER
1	B	51	LYS
1	B	53	THR
1	B	56	ARG
1	B	81	LYS
1	B	83	LEU
1	B	106	LYS
1	B	113	SER
1	B	137	THR
1	B	141	LYS
1	B	147	VAL
1	B	155	LYS
1	B	159	LEU
1	B	184	GLN
1	B	230	GLU
1	B	262	ILE
1	B	269	SER
1	B	308	LEU
1	B	367	SER
1	B	380	ARG
1	B	385	THR
1	B	399	THR
1	B	419	LEU
1	B	431	SER
1	B	506	SER
1	B	520	THR
1	B	552	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	556	SER
1	C	14	CYS
1	C	51	LYS
1	C	53	THR
1	C	56	ARG
1	C	77	THR
1	C	100	LYS
1	C	106	LYS
1	C	113	SER
1	C	130	THR
1	C	139	MET
1	C	155	LYS
1	C	159	LEU
1	C	160	ILE
1	C	184	GLN
1	C	190	SER
1	C	230	GLU
1	C	257	THR
1	C	321	VAL
1	C	336	LEU
1	C	355	GLN
1	C	371	SER
1	C	372	VAL
1	C	380	ARG
1	C	434	LEU
1	C	487	SER
1	C	494	VAL
1	C	506	SER
1	C	539	ILE
1	C	556	SER
1	D	14	CYS
1	D	21	LEU
1	D	31	LEU
1	D	53	THR
1	D	56	ARG
1	D	90	LYS
1	D	100	LYS
1	D	106	LYS
1	D	112	SER
1	D	124	LYS
1	D	137	THR
1	D	148	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	158	ARG
1	D	159	LEU
1	D	207	THR
1	D	226	SER
1	D	230	GLU
1	D	254	LYS
1	D	257	THR
1	D	262	ILE
1	D	267	THR
1	D	272	GLN
1	D	282	SER
1	D	318	ASP
1	D	403	THR
1	D	461	GLN
1	D	492	LEU
1	D	494	VAL
1	D	498	ARG
1	D	539	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (20) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	33	HIS
1	A	58	GLN
1	A	199	GLN
1	A	231	ASN
1	A	330	GLN
1	A	355	GLN
1	A	562	HIS
1	B	33	HIS
1	B	58	GLN
1	B	199	GLN
1	B	355	GLN
1	B	374	HIS
1	B	502	HIS
1	C	199	GLN
1	C	355	GLN
1	C	502	HIS
1	D	33	HIS
1	D	58	GLN
1	D	199	GLN
1	D	273	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)

4 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	3AE	A	601	-	11,13,13	1.11	1 (9%)	16,19,19	1.16	3 (18%)
2	3AE	C	601	-	11,13,13	0.74	1 (9%)	16,19,19	1.62	5 (31%)
2	3AE	B	601	-	11,13,13	0.83	0	16,19,19	1.46	3 (18%)
2	3AE	D	601	-	11,13,13	0.75	1 (9%)	16,19,19	1.41	4 (25%)

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	3AE	A	601	-	-	0/6/10/10	0/1/1/1
2	3AE	C	601	-	-	0/6/10/10	0/1/1/1
2	3AE	B	601	-	-	0/6/10/10	0/1/1/1
2	3AE	D	601	-	-	0/6/10/10	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	3AE	C02-C03	3.30	1.50	1.47
2	C	601	3AE	C02-C03	2.26	1.49	1.47
2	D	601	3AE	C02-C03	2.06	1.49	1.47

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	601	3AE	C06-C02-C03	-2.97	116.38	120.37
2	B	601	3AE	C01-C02-C03	-2.68	116.77	120.37
2	B	601	3AE	C04-C05-C07	2.41	123.80	119.97
2	D	601	3AE	C04-C05-C07	2.40	123.77	119.97
2	C	601	3AE	C01-C02-C03	2.29	123.45	120.37
2	C	601	3AE	C-C05-C07	2.23	123.51	119.97
2	A	601	3AE	C04-C05-C07	2.23	123.51	119.97
2	C	601	3AE	C04-C05-C07	-2.18	116.50	119.97
2	D	601	3AE	F02-C07-C05	-2.17	108.16	112.93
2	C	601	3AE	F01-C07-C05	-2.14	108.22	112.93
2	A	601	3AE	C06-C02-C03	2.10	123.19	120.37
2	B	601	3AE	F01-C07-C05	-2.07	108.38	112.93
2	D	601	3AE	C01-C02-C03	-2.06	117.60	120.37
2	D	601	3AE	C-C05-C07	-2.05	116.71	119.97
2	A	601	3AE	C01-C02-C03	-2.03	117.64	120.37

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 (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	544/566 (96%)	-0.23	2 (0%)	92	93	19, 31, 49, 88	0
1	B	544/566 (96%)	-0.18	6 (1%)	80	82	18, 32, 52, 88	0
1	C	544/566 (96%)	-0.22	2 (0%)	92	93	19, 30, 51, 90	0
1	D	544/566 (96%)	0	100	100		18, 30, 50, 88	0
All	All	2176/2264 (96%)	-0.21	10 (0%)	91	91	18, 31, 51, 90	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	29	SER	2.7
1	B	10	LEU	2.7
1	B	9	ALA	2.6
1	C	7	THR	2.4
1	C	29	SER	2.4
1	B	274	CYS	2.2
1	B	43	ARG	2.2
1	A	31	LEU	2.1
1	B	272	GLN	2.0
1	B	15	ALA	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 monosaccharides 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. 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
2	3AE	C	601	13/13	0.91	0.26	44,52,55,60	0
2	3AE	B	601	13/13	0.91	0.24	54,64,68,68	0
2	3AE	D	601	13/13	0.93	0.18	40,42,45,48	0
2	3AE	A	601	13/13	0.94	0.20	31,36,40,41	0

6.5 Other polymers (i)

There are no such residues in this entry.