WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 10:45 pm BST

PDB ID	:	4A0L
Title	:	Structure of DDB1-DDB2-CUL4B-RBX1 bound to a 12 bp abasic site containing DNA-duplex
Authors	:	Fischer, E.S.; Scrima, A.; Gut, H.; Thoma, N.H.
Deposited on	:	2011-09-09
Resolution	:	7.40 Å(reported)

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

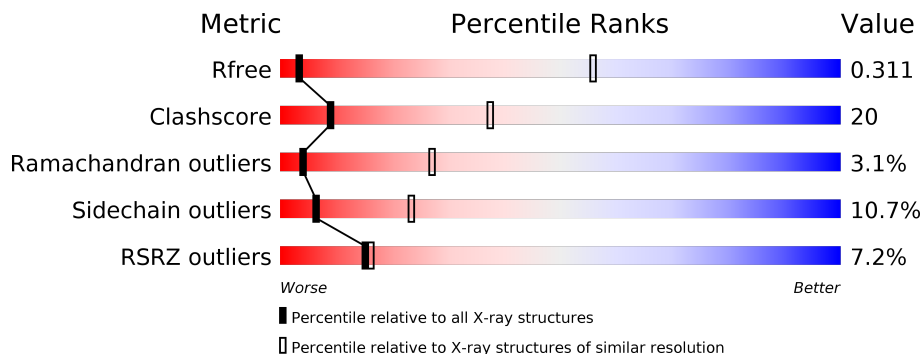|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 7.40 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		0.311
Clashscore		20
Ramachandran outliers		3.1%
Sidechain outliers		10.7%
RSRZ outliers		7.2%

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	1004 (10.00-3.90)
Clashscore	141614	1069 (10.00-3.90)
Ramachandran outliers	138981	1002 (10.00-3.90)
Sidechain outliers	138945	1002 (10.00-3.86)
RSRZ outliers	127900	1004 (9.50-3.80)

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% 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	1144	4% 64% 25% 7% • •
1	C	1144	5% 65% 25% 6% • •
2	B	382	19% 74% 16% • 7%
2	D	382	6% 74% 17% • 7%
3	E	726	6% 59% 32% 6% •
3	H	726	7% 61% 31% 5% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	98	2% / 10% / 7% · · / 79%
4	I	98	8% / 13% / 7% · / 79%
5	R	12	33% / 75% / 25%
5	T	12	17% / 75% / 25%
6	S	12	33% / 50% / 50%
6	U	12	33% / 50% / 50%

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
5	3DR	T	9	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 35553 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 DNA DAMAGE-BINDING PROTEIN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1105	Total 8517	C 5403	N 1417	O 1652	S 45	0	0	0
1	C	1105	Total 8537	C 5409	N 1428	O 1655	S 45	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q16531
A	-2	GLY	-	expression tag	UNP Q16531
A	-1	GLY	-	expression tag	UNP Q16531
A	0	ARG	-	expression tag	UNP Q16531
C	-3	GLY	-	expression tag	UNP Q16531
C	-2	GLY	-	expression tag	UNP Q16531
C	-1	GLY	-	expression tag	UNP Q16531
C	0	ARG	-	expression tag	UNP Q16531

- Molecule 2 is a protein called DNA DAMAGE-BINDING PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	355	Total 2819	C 1792	N 492	O 524	S 11	0	0	0
2	D	355	Total 2843	C 1806	N 499	O 527	S 11	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	76	MET	-	expression tag	UNP Q2YDS1
B	77	HIS	-	expression tag	UNP Q2YDS1
B	78	HIS	-	expression tag	UNP Q2YDS1
B	79	HIS	-	expression tag	UNP Q2YDS1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	80	HIS	-	expression tag	UNP Q2YDS1
B	81	HIS	-	expression tag	UNP Q2YDS1
B	82	HIS	-	expression tag	UNP Q2YDS1
B	83	ARG	-	expression tag	UNP Q2YDS1
B	84	ARG	-	expression tag	UNP Q2YDS1
B	85	LEU	-	expression tag	UNP Q2YDS1
B	86	VAL	-	expression tag	UNP Q2YDS1
B	87	PRO	-	expression tag	UNP Q2YDS1
B	88	ARG	-	expression tag	UNP Q2YDS1
B	89	GLY	-	expression tag	UNP Q2YDS1
B	90	SER	-	expression tag	UNP Q2YDS1
B	91	GLY	-	expression tag	UNP Q2YDS1
B	92	GLY	-	expression tag	UNP Q2YDS1
B	93	ARG	-	expression tag	UNP Q2YDS1
B	180	GLN	LEU	variant	UNP Q2YDS1
B	214	ARG	TRP	variant	UNP Q2YDS1
D	76	MET	-	expression tag	UNP Q2YDS1
D	77	HIS	-	expression tag	UNP Q2YDS1
D	78	HIS	-	expression tag	UNP Q2YDS1
D	79	HIS	-	expression tag	UNP Q2YDS1
D	80	HIS	-	expression tag	UNP Q2YDS1
D	81	HIS	-	expression tag	UNP Q2YDS1
D	82	HIS	-	expression tag	UNP Q2YDS1
D	83	ARG	-	expression tag	UNP Q2YDS1
D	84	ARG	-	expression tag	UNP Q2YDS1
D	85	LEU	-	expression tag	UNP Q2YDS1
D	86	VAL	-	expression tag	UNP Q2YDS1
D	87	PRO	-	expression tag	UNP Q2YDS1
D	88	ARG	-	expression tag	UNP Q2YDS1
D	89	GLY	-	expression tag	UNP Q2YDS1
D	90	SER	-	expression tag	UNP Q2YDS1
D	91	GLY	-	expression tag	UNP Q2YDS1
D	92	GLY	-	expression tag	UNP Q2YDS1
D	93	ARG	-	expression tag	UNP Q2YDS1
D	180	GLN	LEU	variant	UNP Q2YDS1
D	214	ARG	TRP	variant	UNP Q2YDS1

- Molecule 3 is a protein called CULLIN-4B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	E	709	5743	3659	979	1074	31	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	709	Total 5773	C 3681	N 978	O 1082	S 32	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	189	GLY	-	expression tag	UNP Q13620
E	190	GLY	-	expression tag	UNP Q13620
E	191	GLY	-	expression tag	UNP Q13620
E	192	ARG	-	expression tag	UNP Q13620
H	189	GLY	-	expression tag	UNP Q13620
H	190	GLY	-	expression tag	UNP Q13620
H	191	GLY	-	expression tag	UNP Q13620
H	192	ARG	-	expression tag	UNP Q13620

- Molecule 4 is a protein called E3 UBIQUITIN-PROTEIN LIGASE RBX1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	F	21	Total 175	C 118	N 31	O 26	0	0	0
4	I	21	Total 180	C 122	N 32	O 26	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	11	MET	-	expression tag	UNP P62878
I	11	MET	-	expression tag	UNP P62878

- Molecule 5 is a DNA chain called 12 BP THF CONTAINING DNA DUPLEX.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	R	12	Total 234	C 111	N 41	O 70	P 12	0	0	0
5	T	12	Total 234	C 111	N 41	O 70	P 12	0	0	0

- Molecule 6 is a DNA chain called 12 BP DNA DUPLEX.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	S	12	Total 249	C 118	N 47	O 72	P 12	0	0	0

Continued on next page...

Continued from previous page...

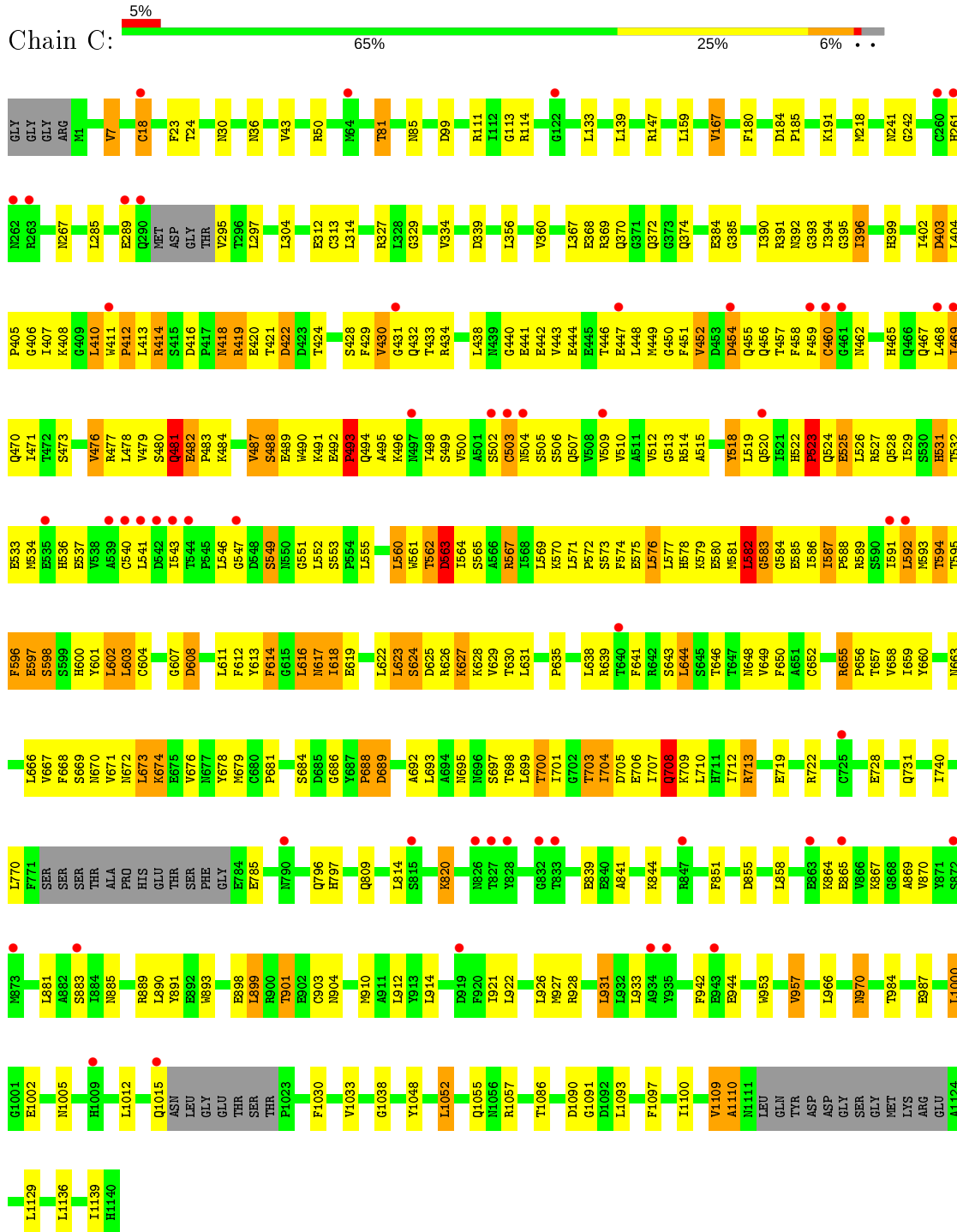| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
6	U	12	249	118	47	72	12	0	0	0

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 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.

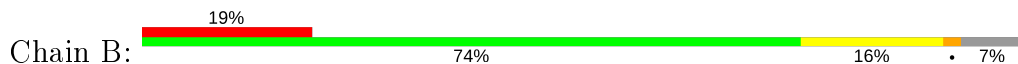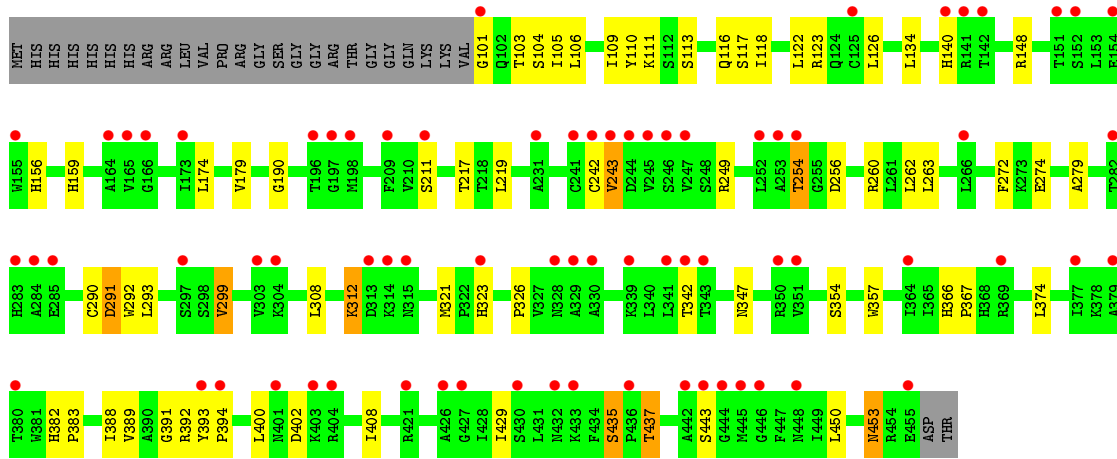• Molecule 1: DNA DAMAGE-BINDING PROTEIN 1

ASP GLY SER MET LYS ARG GLU A1124 L1129 I1130 K1131 L1136 I1139 H1140

● Molecule 1: DNA DAMAGE-BINDING PROTEIN 1

Chain C:

● Molecule 2: DNA DAMAGE-BINDING PROTEIN 2

Chain B:

MET HIS HIS HIS HIS HIS ARG ARG LEU VAL PRO ARG GLY SER GLY ARG THR GLY GLY GLN LYS LYS VAL G101 Q102 T103 S104 I105 L106 I109 Y110 K111 S112 S113 Q116 S117 I118 Q116 S117 I118 L122 Q124 C125 L126 L134 H140 R141 T142 R148 T151 S152 L153 E154

W155 H156 H159 A164 V165 G166 I173 L174 V179 G190 T196 G197 M198 F209 V210 S211 T217 T218 L219 A231 C241 C242 V243 D244 V245 S246 V247 S248 R249 L252 A253 T254 G255 D256 R260 L261 L262 L263 L266 F272 K273 E274 A279 T282

H283 A284 E285 C290 D291 W292 L293 S297 S298 V299 V303 K304 L308 K312 D313 K314 N315 M321 P322 H323 P326 V327 N328 A329 A330 K339 L340 L341 T342 T343 N347 R350 V351 S354 W357 I364 I365 H366 P367 R368 R369 L374 I377 K378 A379

T380 W381 H382 P383 I388 V389 A390 G391 R392 Y393 P394 L400 N401 D402 K403 R404 I408 R421 A426 G427 I428 I429 S430 L431 N432 K433 F434 S435 P436 T437 A442 S443 G444 M445 G446 F447 N448 I449 L450 N453 R454 E455 ASP THR

● **Molecule 2: DNA DAMAGE-BINDING PROTEIN 2**

Chain D:

MET HIS HIS HIS HIS HIS ARG ARG LEU VAL PRO ARG GLY SER GLY ARG THR GLY GLY GLN LYS LYS VAL G101 Q102 T103 S104 I105 L106 I109 Y110 K111 S112 S113 L114 G115 Q116 S117 I118 H119 L126 L134 H140 P145 R148 S152 L153 E154 W155 H156

H159 A164 I173 L174 V179 Q187 G190 T196 G197 M198 S211 T217 T218 L219 V226 C242 V243 R249 T254 G255 D256 R260 L261 L262 L263 F272 K273 E274 A279 H283 C290 D291 W292 L293 V299 L308 K312 M321 P322

H323 P326 T338 T342 N347 S354 Y355 D356 W357 H366 P367 L374 T380 W381 H382 P383 D386 I387 I388 V389 A390 G391 R392 Y393 P394 L400 N401 D402 I408 R421 D422 P423 N424 I429 S430 L431 N432 K433 F434 S435 P436 T437 S443 G444 M445
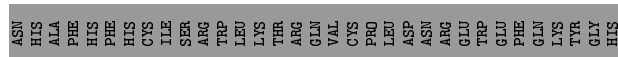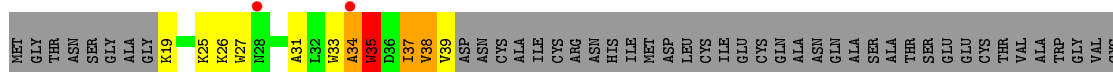
L450 N453 R454 E455 ASP THR

● **Molecule 3: CULLIN-4B**

Chain E:

GLY GLY GLY ARG GLY SER ALA K196 K197 L198 V199 I200 K201 N202 F203 D205 K204 K206 P207 LYS LEU PRO E210 N211 D214 E215 T216 W217 Q218 K219 I227 I233 K234 Y235 N236 L237 E238 Y241 K251 N255 L256 Q259 L260 I263 K269 F275 F285 L286 K287

K288 I289 W293 C297 R298 Q299 M300 I301 M302 I303 L308 F309 L310 D311 V315 L316 I324 F332 H335 I336 I337 D338 D339 Q340 K341 V342 Q343 N344 K345 D348 G349 I350 L351 L352 L353 I354 E355 R356 R357 R358 N359 I363 D364 R365 S366 L367 L368 L372 S373 M374

L375 S376 D377 L378 Q379 I380 S384 L390 E391 R395 L396 Y397 L404 R408 E409 V410 P411 L414 H415 H416 V417 N418 K419 R420 D426 R427 L428 I429 T430 Y431 L432 D433 Q434 T435 T436 Q437 L440 T443 V444 E445 L448 L449 G450 E451 H452 L453 T454 Q458 K459

G460 L464 L465 D466 E467 M468 R469 I470 I473 S474 L475 L479 R484 G485 G486 V487 Q488 L491 Q492 Q493 W494 F501 G502 S503 T504 I505 V506 I507 N508 P509 E510 K511 D512 K513 T514 M515 V516 L520 K523 D524 I530 D531 I532 C533 F534 N537 E538 K539 F540 I541

● Molecule 3: CULLIN-4B

● Molecule 4: E3 UBIQUITIN-PROTEIN LIGASE RBX1

Chain F:

2%
10% 7% 79%

MET GLY THR ASN SER GLY ALA GLY K19 K25 K26 W27 N28 A31 L32 W33 A34 K35 D36 I37 V38 V39 ASP ASN CYS ALA ILE CYS ARG ASN HIS ILE MET ASP LEU CYS ILE GLU CYS GLN ALA ASN GLN ALA SER ALA THR SER GLU GLU CYS THR VAL ALA TRP GLY VAL CYS

ASN HIS ALA PHE HIS PHE HIS CYS ILE SER ARG TRP LEU LYS THR ARG GLN VAL CYS PRO LEU ASP ASN ARG GLU TRP GLU GLN PHE GLN LYS TYR GLY HIS

● Molecule 4: E3 UBIQUITIN-PROTEIN LIGASE RBX1

Chain I:

8%
13% 7% 79%

MET GLY THR ASN SER GLY ALA GLY K19 K20 R21 K25 K26 W27 N28 A31 L32 W33 A34 W35 D36 I37 V38 V39 ASP ASN CYS ALA ILE CYS ARG ASN HIS ILE MET ASP LEU CYS ILE GLU CYS GLN ALA ASN GLN ALA SER ALA THR SER GLU GLU CYS THR VAL ALA TRP GLY

VAL CYS ASN HIS ALA PHE HIS PHE HIS CYS ILE SER ARG TRP LEU LYS THR ARG GLN VAL CYS PRO LEU ASP ASN ARG GLU TRP GLU PHE GLN LYS TYR GLY HIS

● Molecule 5: 12 BP THF CONTAINING DNA DUPLEX

Chain R:

33%
75% 25%

G3 C4 T5 A6 C11 G12 C13 A14

● Molecule 5: 12 BP THF CONTAINING DNA DUPLEX

Chain T:

17%
75% 25%

G3 A6 A10 C11 G12 C13 A14

● Molecule 6: 12 BP DNA DUPLEX

Chain S:

33%
50% 50%

T3 G4 C5 G6 T7 A8 A9 A12 G13 C14

● Molecule 6: 12 BP DNA DUPLEX

Chain U:

33%
50% 50%

T3 G4 C5 G6 T7 A8 A9 G10 T11 A12 G13 C14

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	130.80Å 155.84Å 255.39Å 90.00° 94.17° 90.00°	Depositor
Resolution (Å)	29.86 – 7.40 29.86 – 7.40	Depositor EDS
% Data completeness (in resolution range)	99.0 (29.86-7.40) 99.2 (29.86-7.40)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.74 (at 7.23Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.318 , 0.320 0.313 , 0.311	Depositor DCC
R_{free} test set	1355 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	444.7	Xtriage
Anisotropy	0.287	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 213.2	EDS
L-test for twinning[2]	$<|L|> = 0.45$, $<L^2> = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.76	EDS
Total number of atoms	35553	wwPDB-VP
Average B, all atoms (Å^2)	253.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.82% 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: 3DR

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	0.43	2/8668 (0.0%)	0.66	7/11756 (0.1%)
1	C	0.43	2/8688 (0.0%)	0.66	8/11783 (0.1%)
2	B	0.38	0/2891	0.60	0/3928
2	D	0.38	0/2917	0.60	0/3962
3	E	0.40	3/5831 (0.1%)	0.55	1/7832 (0.0%)
3	H	0.38	0/5865	0.54	0/7878
4	F	0.35	0/179	0.39	0/241
4	I	0.42	0/186	0.43	0/251
5	R	0.92	1/248 (0.4%)	1.13	0/377
5	T	0.92	1/248 (0.4%)	1.12	0/377
6	S	0.89	0/279	1.36	1/429 (0.2%)
6	U	0.88	0/279	1.36	1/429 (0.2%)
All	All	0.43	9/36279 (0.0%)	0.64	18/49243 (0.0%)

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	582	LEU	C-N	6.82	1.45	1.33
1	A	582	LEU	C-N	6.78	1.45	1.33
3	E	579	THR	C-N	-5.89	1.20	1.34
3	E	376	SER	C-N	5.80	1.47	1.34
5	R	11	DC	C1'-N1	5.47	1.56	1.49

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	624	SER	N-CA-C	7.04	130.01	111.00
1	C	624	SER	N-CA-C	7.02	129.95	111.00
1	A	688	PRO	N-CA-C	6.63	129.34	112.10
1	C	688	PRO	N-CA-C	6.56	129.16	112.10
1	A	689	ASP	N-CA-C	-6.55	93.32	111.00

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	8517	0	8394	471	0
1	C	8537	0	8432	389	0
2	B	2819	0	2745	88	0
2	D	2843	0	2788	74	0
3	E	5743	0	5773	296	0
3	H	5773	0	5798	241	0
4	F	175	0	184	22	0
4	I	180	0	178	6	0
5	R	234	0	132	3	0
5	T	234	0	132	3	0
6	S	249	0	136	7	0
6	U	249	0	136	7	0
All	All	35553	0	34828	1424	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 20.

The worst 5 of 1424 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)
3:E:555:ARG:CG	3:E:559:PRO:HD3	1.32	1.56
1:A:354:THR:CG2	1:A:712:ILE:HG21	1.37	1.53
3:E:552:ILE:CG2	3:E:597:ILE:HG12	1.42	1.48
3:H:365:ARG:CB	3:H:430:THR:HG22	1.44	1.45
3:E:555:ARG:HB3	3:E:559:PRO:CG	1.52	1.37

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	1095/1144 (96%)	948 (87%)	105 (10%)	42 (4%)	3	24
1	C	1095/1144 (96%)	967 (88%)	97 (9%)	31 (3%)	5	30
2	B	353/382 (92%)	333 (94%)	16 (4%)	4 (1%)	14	52
2	D	353/382 (92%)	333 (94%)	16 (4%)	4 (1%)	14	52
3	E	701/726 (97%)	603 (86%)	73 (10%)	25 (4%)	3	25
3	H	701/726 (97%)	604 (86%)	72 (10%)	25 (4%)	3	25
4	F	19/98 (19%)	14 (74%)	2 (10%)	3 (16%)	0	3
4	I	19/98 (19%)	16 (84%)	2 (10%)	1 (5%)	2	19
All	All	4336/4700 (92%)	3818 (88%)	383 (9%)	135 (3%)	4	27

5 of 135 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	367	LEU
1	A	390	ILE
1	A	430	VAL
1	A	583	GLY
1	A	598	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	938/1000 (94%)	847 (90%)	91 (10%)	8	27

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	944/1000 (94%)	854 (90%)	90 (10%)	8	27
2	B	306/335 (91%)	283 (92%)	23 (8%)	13	38
2	D	313/335 (93%)	290 (93%)	23 (7%)	14	39
3	E	625/660 (95%)	535 (86%)	90 (14%)	3	16
3	H	631/660 (96%)	548 (87%)	83 (13%)	4	18
4	F	17/83 (20%)	13 (76%)	4 (24%)	1	4
4	I	16/83 (19%)	14 (88%)	2 (12%)	4	19
All	All	3790/4156 (91%)	3384 (89%)	406 (11%)	6	23

5 of 406 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	728	GLU
3	E	198	LEU
3	H	595	ARG
1	C	881	LEU
2	D	134	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 99 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	494	GLN
1	C	885	ASN
3	H	557	ASN
1	C	524	GLN
1	C	648	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)

2 non-standard protein/DNA/RNA residues 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	3DR	R	9	5	8,11,12	0.40	0	9,14,17	1.10	0
5	3DR	T	9	5	8,11,12	0.41	0	9,14,17	1.09	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	3DR	R	9	5	-	1/3/15/16	0/1/1/1
5	3DR	T	9	5	-	1/3/15/16	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	R	9	3DR	C3'-C4'-C5'-O5'
5	T	9	3DR	C3'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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	1105/1144 (96%)	0.42	49 (4%)	34	32	98, 183, 251, 326	0
1	C	1105/1144 (96%)	0.49	55 (4%)	28	28	130, 197, 349, 378	0
2	B	355/382 (92%)	1.05	74 (20%)	1	3	275, 328, 394, 423	0
2	D	355/382 (92%)	0.66	23 (6%)	18	18	193, 275, 343, 370	0
3	E	709/726 (97%)	0.44	44 (6%)	20	20	131, 321, 394, 414	0
3	H	709/726 (97%)	0.49	49 (6%)	16	17	226, 289, 381, 399	0
4	F	21/98 (21%)	0.78	2 (9%)	8	11	368, 383, 396, 398	0
4	I	21/98 (21%)	1.53	8 (38%)	0	1	340, 358, 380, 389	0
5	R	11/12 (91%)	1.86	4 (36%)	0	1	301, 317, 376, 403	0
5	T	11/12 (91%)	1.32	2 (18%)	1	4	257, 344, 355, 356	0
6	S	12/12 (100%)	1.74	4 (33%)	0	2	317, 326, 386, 386	0
6	U	12/12 (100%)	1.69	4 (33%)	0	2	286, 320, 350, 363	0
All	All	4426/4748 (93%)	0.54	318 (7%)	15	16	98, 240, 379, 423	0

The worst 5 of 318 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	E	775	ASP	10.5
1	C	542	ASP	7.1
3	E	810	CYS	6.3
1	C	503	CYS	5.8
2	D	430	SER	5.8

6.2 Non-standard residues in protein, DNA, RNA chains (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
5	3DR	T	9	11/12	0.69	0.68	347,358,362,362	0
5	3DR	R	9	11/12	0.78	0.38	318,324,336,339	0

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.