W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2021 – 07:06 AM EDT

PDB ID	:	2P5O
Title	:	Crystal structure of RB69 GP43 in complex with DNA containing an abasic site analog
Authors	:	Hogg, M.; Wallace, S.S.; Doublie, S.
Deposited on	:	2007-03-15
Resolution	:	2.80 Å(reported)

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

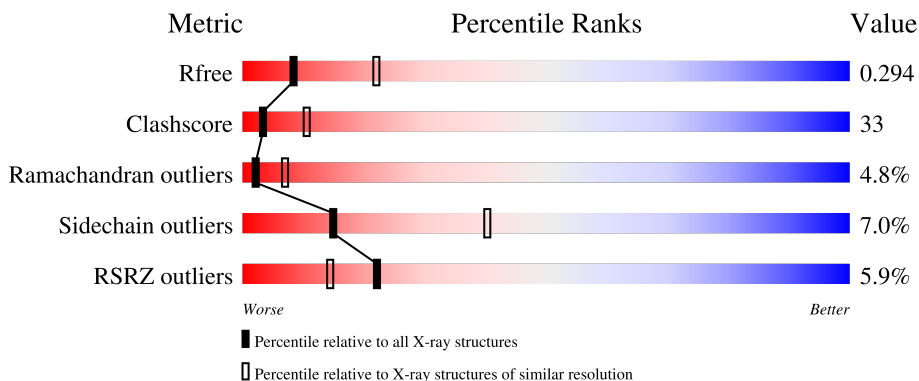|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.80 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		0.294
Clashscore		33
Ramachandran outliers		4.8%
Sidechain outliers		7.0%
RSRZ outliers		5.9%

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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	E	18	6% — 44% / 56%
1	G	18	6% — 17% / 56% / 28%
1	I	18	44% / 50% / 6%
1	K	18	22% — 6% / 44% / 50%
2	F	15	7% — 33% / 67%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	15	27% / 40% / 60%
2	J	15	7% / 33% / 67%
2	L	15	47% / 7% / 47% / 47%
3	A	903	% / 56% / 35% / 5%
3	B	903	5% / 39% / 40% / 5% / 15%
3	C	903	% / 51% / 38% / 6% / 6%
3	D	903	11% / 30% / 37% / 7% / 26%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 27752 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 DNA chain called Template DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	18	Total 355	C 169	N 64	O 105	P 17	0	0	0
1	G	13	Total 264	C 126	N 51	O 75	P 12	0	0	0
1	I	18	Total 355	C 169	N 64	O 105	P 17	0	0	0
1	K	9	Total 181	C 86	N 37	O 50	P 8	0	0	0

- Molecule 2 is a DNA chain called Primer DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	15	Total 308	C 147	N 60	O 87	P 14	0	0	0
2	H	15	Total 308	C 147	N 60	O 87	P 14	0	0	0
2	J	15	Total 308	C 147	N 60	O 87	P 14	0	0	0
2	L	8	Total 163	C 78	N 30	O 48	P 7	0	0	0

- Molecule 3 is a protein called DNA polymerase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	A	856	Total 6941	C 4460	N 1154	O 1295	S 8	Se 24	0	0	0
3	B	771	Total 6245	C 4013	N 1034	O 1167	S 6	Se 25	0	0	0
3	C	853	Total 6895	C 4430	N 1143	O 1290	S 8	Se 24	0	0	0
3	D	671	Total 4934	C 3144	N 818	O 948	S 6	Se 18	0	0	0

There are 108 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	modified residue	UNP Q38087
A	65	MSE	MET	modified residue	UNP Q38087
A	75	MSE	MET	modified residue	UNP Q38087
A	85	MSE	MET	modified residue	UNP Q38087
A	189	MSE	MET	modified residue	UNP Q38087
A	199	MSE	MET	modified residue	UNP Q38087
A	222	ALA	ASP	engineered mutation	UNP Q38087
A	256	MSE	MET	modified residue	UNP Q38087
A	327	ALA	ASP	engineered mutation	UNP Q38087
A	347	MSE	MET	modified residue	UNP Q38087
A	408	MSE	MET	modified residue	UNP Q38087
A	461	MSE	MET	modified residue	UNP Q38087
A	462	MSE	MET	modified residue	UNP Q38087
A	489	MSE	MET	modified residue	UNP Q38087
A	541	MSE	MET	modified residue	UNP Q38087
A	553	MSE	MET	modified residue	UNP Q38087
A	592	MSE	MET	modified residue	UNP Q38087
A	659	MSE	MET	modified residue	UNP Q38087
A	670	MSE	MET	modified residue	UNP Q38087
A	674	MSE	MET	modified residue	UNP Q38087
A	681	MSE	MET	modified residue	UNP Q38087
A	683	MSE	MET	modified residue	UNP Q38087
A	715	MSE	MET	modified residue	UNP Q38087
A	728	MSE	MET	modified residue	UNP Q38087
A	752	MSE	MET	modified residue	UNP Q38087
A	866	MSE	MET	modified residue	UNP Q38087
A	900	MSE	MET	modified residue	UNP Q38087
B	1	MSE	-	modified residue	UNP Q38087
B	65	MSE	MET	modified residue	UNP Q38087
B	75	MSE	MET	modified residue	UNP Q38087
B	85	MSE	MET	modified residue	UNP Q38087
B	189	MSE	MET	modified residue	UNP Q38087
B	199	MSE	MET	modified residue	UNP Q38087
B	222	ALA	ASP	engineered mutation	UNP Q38087
B	256	MSE	MET	modified residue	UNP Q38087
B	327	ALA	ASP	engineered mutation	UNP Q38087
B	347	MSE	MET	modified residue	UNP Q38087
B	408	MSE	MET	modified residue	UNP Q38087
B	461	MSE	MET	modified residue	UNP Q38087
B	462	MSE	MET	modified residue	UNP Q38087
B	489	MSE	MET	modified residue	UNP Q38087
B	541	MSE	MET	modified residue	UNP Q38087

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	553	MSE	MET	modified residue	UNP Q38087
B	592	MSE	MET	modified residue	UNP Q38087
B	659	MSE	MET	modified residue	UNP Q38087
B	670	MSE	MET	modified residue	UNP Q38087
B	674	MSE	MET	modified residue	UNP Q38087
B	681	MSE	MET	modified residue	UNP Q38087
B	683	MSE	MET	modified residue	UNP Q38087
B	715	MSE	MET	modified residue	UNP Q38087
B	728	MSE	MET	modified residue	UNP Q38087
B	752	MSE	MET	modified residue	UNP Q38087
B	866	MSE	MET	modified residue	UNP Q38087
B	900	MSE	MET	modified residue	UNP Q38087
C	1	MSE	-	modified residue	UNP Q38087
C	65	MSE	MET	modified residue	UNP Q38087
C	75	MSE	MET	modified residue	UNP Q38087
C	85	MSE	MET	modified residue	UNP Q38087
C	189	MSE	MET	modified residue	UNP Q38087
C	199	MSE	MET	modified residue	UNP Q38087
C	222	ALA	ASP	engineered mutation	UNP Q38087
C	256	MSE	MET	modified residue	UNP Q38087
C	327	ALA	ASP	engineered mutation	UNP Q38087
C	347	MSE	MET	modified residue	UNP Q38087
C	408	MSE	MET	modified residue	UNP Q38087
C	461	MSE	MET	modified residue	UNP Q38087
C	462	MSE	MET	modified residue	UNP Q38087
C	489	MSE	MET	modified residue	UNP Q38087
C	541	MSE	MET	modified residue	UNP Q38087
C	553	MSE	MET	modified residue	UNP Q38087
C	592	MSE	MET	modified residue	UNP Q38087
C	659	MSE	MET	modified residue	UNP Q38087
C	670	MSE	MET	modified residue	UNP Q38087
C	674	MSE	MET	modified residue	UNP Q38087
C	681	MSE	MET	modified residue	UNP Q38087
C	683	MSE	MET	modified residue	UNP Q38087
C	715	MSE	MET	modified residue	UNP Q38087
C	728	MSE	MET	modified residue	UNP Q38087
C	752	MSE	MET	modified residue	UNP Q38087
C	866	MSE	MET	modified residue	UNP Q38087
C	900	MSE	MET	modified residue	UNP Q38087
D	1	MSE	-	modified residue	UNP Q38087
D	65	MSE	MET	modified residue	UNP Q38087
D	75	MSE	MET	modified residue	UNP Q38087

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	85	MSE	MET	modified residue	UNP Q38087
D	189	MSE	MET	modified residue	UNP Q38087
D	199	MSE	MET	modified residue	UNP Q38087
D	222	ALA	ASP	engineered mutation	UNP Q38087
D	256	MSE	MET	modified residue	UNP Q38087
D	327	ALA	ASP	engineered mutation	UNP Q38087
D	347	MSE	MET	modified residue	UNP Q38087
D	408	MSE	MET	modified residue	UNP Q38087
D	461	MSE	MET	modified residue	UNP Q38087
D	462	MSE	MET	modified residue	UNP Q38087
D	489	MSE	MET	modified residue	UNP Q38087
D	541	MSE	MET	modified residue	UNP Q38087
D	553	MSE	MET	modified residue	UNP Q38087
D	592	MSE	MET	modified residue	UNP Q38087
D	659	MSE	MET	modified residue	UNP Q38087
D	670	MSE	MET	modified residue	UNP Q38087
D	674	MSE	MET	modified residue	UNP Q38087
D	681	MSE	MET	modified residue	UNP Q38087
D	683	MSE	MET	modified residue	UNP Q38087
D	715	MSE	MET	modified residue	UNP Q38087
D	728	MSE	MET	modified residue	UNP Q38087
D	752	MSE	MET	modified residue	UNP Q38087
D	866	MSE	MET	modified residue	UNP Q38087
D	900	MSE	MET	modified residue	UNP Q38087

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	E	3	Total 3	O 3	0	0
4	F	5	Total 5	O 5	0	0
4	G	7	Total 7	O 7	0	0
4	H	2	Total 2	O 2	0	0
4	I	17	Total 17	O 17	0	0
4	J	9	Total 9	O 9	0	0
4	K	8	Total 8	O 8	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	L	6	Total 6	O 6	0	0
4	A	139	Total 139	O 139	0	0
4	B	106	Total 106	O 106	0	0
4	C	147	Total 147	O 147	0	0
4	D	46	Total 46	O 46	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Template DNA

Chain E:

- Molecule 1: Template DNA

Chain G:

- Molecule 1: Template DNA

Chain I:

- Molecule 1: Template DNA

Chain K:

- Molecule 2: Primer DNA

Chain F:

- Molecule 2: Primer DNA

Chain H:

27%

40%　　　　　　　　　　　　　　　　60%

G101 G104 C105 T106 G107 T108 C109 A113 G114 A115

• Molecule 2: Primer DNA

Chain J:
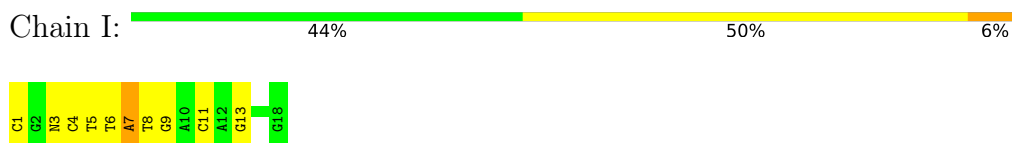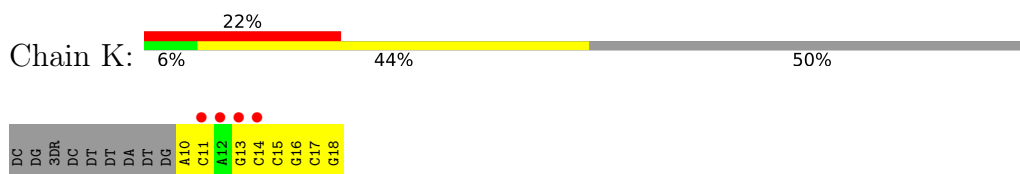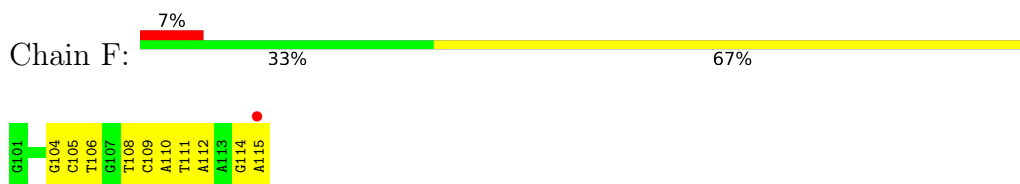
7%

33%　　　　　　　　　　　　　67%

G101 C102 G103 G104 C105 T106 A112 A113 G114 A115

• Molecule 2: Primer DNA

Chain L:

47%

7%　　　　　47%　　　　　　　47%

G101 C102 G103 G104 C105 T106 G107 T108 DC DA DT DA DA DG DA

• Molecule 3: DNA polymerase

Chain A:

%

56%　　　　　　　35%　　　　• 5%

M1 L6 E9 Q10 I11 G12 D13 F16 E17 R18 Y19 I20 G24 R29 E32 Y33 K34 P35 S36 L37 F38 C41 P42 E43 S44 Q45 D51 I52 Y53 G54 K55 K60 A63 N64 M65 R66 D67 W71 I72 M75 E76 L80 E81 A82 L83 G84 M85

K89 L90 A91 Y92 Y97 N98 Y99 D104 H105 T106 K107 I108 N112 F113 D114 I115 E116 V117 D121 G122 K130 H131 P132 A135 I136 T137 D143 L152 N153 S154 V159 W162 S163 E165 I166 A167 A168 K169 L170 Q171 D176 E177 V178 E181 I182

Y188 M189 P190 F191 E194 L197 L198 M199 E200 Y201 L202 N203 F204 W205 Q206 Q207 T214 G215 W216 N217 S220 F221 A222 Y225 V226 K231 E236 A239 K240 R241 L242 I253 E254 S259 R260 E261 S269 V270 L271 D272 Y273 I274 Y277 K278 K279 F280 S281

F282 T283 N284 Q285 Y288 S289 L290 S294 L298 N299 K302 L303 K304 P308 I309 S310 K311 E314 S315 N316 H317 Q318 R319 Y323 I326 A327 V328 Y329 R330 Q333 Q339 F340 I341 N342 M347 I353 Q354 I355 Q356 S357 V358 F359 S360 P361 I362 K363 T364 W365

D366 I369 V379 I380 R384 Y391 A394 F395 V396 K397 E398 N402 R403 Y404 K405 M408 S409 F410 D411 S414 L415 Y416 P417 S418 I419 Q422 V423 N424 I425 S426 P427 E428 T429 P438 L439 H440 D441 Y442 I443 N444 A447 P458 M461 M462 Y463 D466

V470 V471 P472 T473 E474 I475 R482 H485 K486 G487 Y488 M489 L490 A491 A492 Q493 R494 ASN GLY GLU ILE ILE LYS GLU ALA LEU HIS ASN PRO ASN LEU SER VAL ASP GLU PRO LEU ASP VAL ASP TYR ARG PHE ASP PHE SER ASP GLU ILE LYS GLU LYS ILE LYS LEU SER

ALA LYS SER LEU ASN GLU M541 L542 F543 R544 A545 Q546 R547 T548 E549 V550 A551 G552 M553 T554 A555 Q556 I557 N558 R559 K560 L566 N572 R576 R581 N582 A583 Q591 M592 Q595 W596 I597 E598 G610 V617 L618 E619 G620 D621 I622 I625 V626 V627 S628 A629

I632 I633 D634 K635 V636 G637 E638 F641 R642 D643 T644 N645 H646 L651 A655 R656 M659 E660 P661 D664 R668 E669 M670 N676 K677 Q678 H679 I680 M681 F682 D683 D684 R685 L693 G697 I698 G699 G700 F701 W702 K705 Y708 Y720 K724 L725 M728

• Molecule 3: DNA polymerase

GLN VAL VAL GLY GLU LYS VAL TYR VAL LEU PRO LEU ARG GLU GLY ASN PRO PHE GLY ASP LYS CYS ILE ALA TRP PRO SER GLY THR GLU ILE THR ASP LEU ILE LYS ASP D861 V862 L863 H864 W865 M866 D867 V870 L871 L872 T875 F876 I877 K878 P879 F883 L889

K893 K894 A895 S896 L897 F898 M900 F901 ASP PHE

• **Molecule 3: DNA polymerase**

Chain C:

M1 K2 E3 T7 V8 E9 Q10 I11 G12 D13 S14 I15 F16 E17 R18 Y19 I20 D21 S22 N23 G24 R25 E26 R27 L37 F38 P42 E43 S44 Q45 A46 T47 K48 Y49 F50 D51 I52 Y53 G54 K55 T58 R59 K60 L61 F62 M65 Q70 M75 I78 G79 L80 E81

A82 I83 G84 M85 K89 L90 N98 D104 H105 T106 R109 V110 A111 N112 F113 E116 V117 T118 S119 P120 S127 Q128 A129 K130 H131 P132 I133 D134 A135 T137 H138 Y139 D140 S141 I142 D143 D144 R145 F146 Y147 V148 F149 D150 L151 L152 N153 S154 P155 Y156 G157 N158 V159 E160

E161 W162 S163 I164 A167 L170 Q171 E172 Q173 G174 G175 D176 E177 V178 P179 S180 E181 I182 I183 D184 K185 I186 I187 Y188 M189 P190 F191 D192 N193 E194 K195 E196 L197 L198 M199 E200 Y201 L202 N203 F204 W205 Q206 Q207 K208 I212 L213 T214 G215 W216 N217 V218 F221 A222 I223 P224 Y225

V226 Y227 N228 R229 I230 K231 N232 I233 F234 G235 E236 P244 H245 R246 V252 I253 E254 N255 M256 Y257 G258 S259 R260 T264 L265 F266 G267 Y273 L276 Y277 K278 K279 F280 S281 N284 Q285 L290 D291 Y292 I293 S294 E297 L298 N299 Y300 G301 K302 L303 D306 G307

R313 E314 S315 N316 H317 Q318 R319 Y320 I321 S322 Y323 N324 I325 R330 V331 L332 R338 Q339 F340 L343 M347 I353 Q354 I355 Q356 S357 V358 F359 I362 K363 T364 W365 I368 I369 F370 L373 K374 K378 V379 I380 P381 R384 Y391 P392 G393 A394 F395 P401

M402 R403 Y404 K405 M408 S409 F410 D411 L412 T413 S414 L415 Y416 P417 Q422 V423 N424 I425 S426 P427 E428 T433 F434 K435 V436 A437 P438 L439 H440 D441 Y442 I443 A447 E448 R449 P450 S451 D452 V453 Y454 S455 C456 S457 P458 N459 G460 M461 M462 Y463 D466 G469 V470 V471 P472

T473 E474 I475 V478 F479 R482 K486 M489 L490 A491 A492 Q493 R494 N495 GLY GLU ILE ILE LYS GLU VAL ALA LEU HIS ASN PRO ASN LEU SER VAL ASP GLU GLU PRO LEU ASP VAL ASP TYR ARG PHE ASP PHE SER ASP GLU ILE LYS GLU LYS ILE LYS LYS LEU SER ALA LYS SER

LEU ASN GLU M541 L542 F543 R544 E549 V550 A551 G552 M553 T554 A555 Q556 I557 N558 R559 K560 L561 L562 I563 N564 S565 L566 F575 R576 D579 N582 Q591 M592 A593 L594 Q595 L605 N606 E607 V608 C609 G610 L618 T622 D623 S624 A629 D630 K631 I632 I633 V636

W647 F650 L651 E660 P661 A662 I663 D664 R665 G666 F667 R668 E669 M670 M674 N675 N676 K677 Q678 H679 L680 M681 F682 M683 D684 R685 I688 A689 G690 P691 P692 S695 K696 G697 I698 G699 G700 F701 W702 K705 L710 Y720 A721 E722 P723 K724 L725 K726 I727 M728

E731 T732 Q733 K734 P738 V741 Q742 L745 K746 E747 R750 R751 M752 L753 Q754 E755 G756 E757 L760 Y763 E766 E770 F771 R772 Y776 I779 N787 K790 Y791 G795 K800 C801 P802 R813 A814 I815 K816 G817 N818 I819 D820 A821 V825

E828 K829 V830 Y831 V832 L833 A847 W848 I854 L857 L863 M866 L871 L872 E873 K874 T875 F876 I877 K878 P879 L880 E881 A886 L889 E892 K893 K894 L897 F898 ASP MSE PHE ASP PHE

• **Molecule 3: DNA polymerase**

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	133.02Å 123.21Å 165.62Å 90.00° 95.82° 90.00°	Depositor
Resolution (Å)	44.00 – 2.80 44.11 – 2.80	Depositor EDS
% Data completeness (in resolution range)	90.9 (44.00-2.80) 95.8 (44.11-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.39 (at 2.81Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.245 , 0.290 0.253 , 0.294	Depositor DCC
R_{free} test set	24426 reflections (9.67%)	wwPDB-VP
Wilson B-factor (Å^2)	49.8	Xtriage
Anisotropy	0.156	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 60.7	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	27752	wwPDB-VP
Average B, all atoms (Å^2)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.47% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 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	E	0.36	0/384	0.68	0/588
1	G	0.40	0/296	0.91	1/455 (0.2%)
1	I	0.55	0/384	0.78	0/588
1	K	0.37	0/203	0.66	0/311
2	F	0.31	0/346	0.67	0/533
2	H	0.33	0/346	0.68	0/533
2	J	0.43	0/346	0.73	0/533
2	L	0.26	0/182	0.63	0/280
3	A	0.49	0/7090	0.69	0/9555
3	B	0.42	0/6376	0.63	0/8593
3	C	0.46	0/7045	0.66	0/9502
3	D	0.33	0/5028	0.55	0/6831
All	All	0.43	0/28026	0.65	1/38302 (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	G	0	1
1	I	0	1
3	B	0	1
All	All	0	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	6	DT	C6-C5-C7	-5.18	119.79	122.90

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	273	TYR	Sidechain
1	G	17	DC	Sidechain
1	I	7	DA	Sidechain

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	355	0	200	18	0
1	G	264	0	147	12	0
1	I	355	0	200	13	0
1	K	181	0	101	15	0
2	F	308	0	170	20	0
2	H	308	0	170	18	0
2	J	308	0	170	13	0
2	L	163	0	92	24	0
3	A	6941	0	6787	309	0
3	B	6245	0	6024	495	0
3	C	6895	0	6710	380	0
3	D	4934	0	4297	420	0
4	A	139	0	0	18	0
4	B	106	0	0	24	0
4	C	147	0	0	16	0
4	D	46	0	0	9	0
4	E	3	0	0	1	0
4	F	5	0	0	0	0
4	G	7	0	0	0	0
4	H	2	0	0	0	0
4	I	17	0	0	0	0
4	J	9	0	0	0	0
4	K	8	0	0	1	0
4	L	6	0	0	2	0
All	All	27752	0	25068	1711	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 33.

The worst 5 of 1711 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:L:107:DG:C2'	2:L:108:DT:H71	1.77	1.14
1:G:12:DA:H2"	1:G:13:DG:H5'	1.30	1.14
3:D:14:SER:HA	3:D:32:GLU:HA	1.32	1.11
3:B:164:ILE:HD12	3:B:164:ILE:H	1.10	1.10
1:E:16:DG:H2"	1:E:17:DC:H5'	1.27	1.09

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	A	852/903 (94%)	760 (89%)	76 (9%)	16 (2%)	8	26
3	B	763/903 (84%)	621 (81%)	110 (14%)	32 (4%)	3	9
3	C	849/903 (94%)	745 (88%)	74 (9%)	30 (4%)	3	12
3	D	661/903 (73%)	433 (66%)	155 (23%)	73 (11%)	0	1
All	All	3125/3612 (86%)	2559 (82%)	415 (13%)	151 (5%)	2	7

5 of 151 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	12	GLY
3	A	542	LEU
3	A	859	LYS
3	A	894	LYS
3	B	164	ILE

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	
3	A	745/775 (96%)	701 (94%)	44 (6%)	19	49
3	B	663/775 (86%)	613 (92%)	50 (8%)	13	37
3	C	737/775 (95%)	689 (94%)	48 (6%)	17	44
3	D	453/775 (58%)	413 (91%)	40 (9%)	10	29
All	All	2598/3100 (84%)	2416 (93%)	182 (7%)	15	40

5 of 182 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	291	ASP
3	D	17	GLU
3	C	384	ARG
3	C	494	ARG
3	D	92	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 66 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	679	HIS
3	D	70	GLN
3	D	742	GLN
3	B	424	ASN
3	B	389	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)

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	#$|Z| > 2$
1	3DR	E	3	1	8,11,12	0.34	0	9,14,17	0.72	0
1	3DR	I	3	1	8,11,12	0.35	0	9,14,17	0.75	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
1	3DR	E	3	1	-	1/3/15/16	0/1/1/1
1	3DR	I	3	1	-	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
1	I	3	3DR	O4'-C4'-C5'-O5'
1	E	3	3DR	O4'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	3	3DR	3	0
1	I	3	3DR	1	0

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data (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	E	17/18 (94%)	0.57	1 (5%)	22	14	57, 95, 121, 134	0
1	G	13/18 (72%)	1.28	1 (7%)	13	7	48, 103, 143, 144	0
1	I	17/18 (94%)	-0.12	0	100	100	27, 43, 112, 123	0
1	K	9/18 (50%)	1.80	4 (44%)	0	0	38, 140, 155, 159	0
2	F	15/15 (100%)	0.73	1 (6%)	17	10	77, 104, 136, 140	0
2	H	15/15 (100%)	1.31	4 (26%)	0	0	75, 118, 131, 139	0
2	J	15/15 (100%)	0.10	1 (6%)	17	10	28, 57, 97, 114	0
2	L	8/15 (53%)	2.62	7 (87%)	0	0	144, 148, 149, 149	0
3	A	831/903 (92%)	-0.06	9 (1%)	80	75	16, 40, 83, 133	0
3	B	746/903 (82%)	0.29	44 (5%)	22	14	19, 57, 112, 140	0
3	C	829/903 (91%)	0.01	12 (1%)	75	70	18, 46, 86, 119	0
3	D	650/903 (71%)	0.90	102 (15%)	2	1	78, 102, 134, 142	0
All	All	3165/3744 (84%)	0.27	186 (5%)	22	14	16, 56, 123, 159	0

The worst 5 of 186 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	450	PRO	7.1
3	D	309	ILE	6.3
3	D	394	ALA	5.8
3	B	254	GLU	5.5
3	D	216	TRP	5.2

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
1	3DR	E	3	11/12	0.78	0.39	117,125,130,133	0
1	3DR	I	3	11/12	0.89	0.25	101,102,103,103	0

6.3 Carbohydrates (i)

There are no monosaccharides 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.