



# Full wwPDB X-ray Structure Validation Report

Nov 6, 2014 – 06:04 PM EST

PDB ID : 4L2O  
Title : Crystal structure of human ALDH3A1 with its selective inhibitor 1-(4-fluorophenyl)sulfonyl-2-methylbenzimidazole  
Authors : Hurley, T.D.; Parajuli, B.  
Deposited on : 2013-06-04  
Resolution : 1.94 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

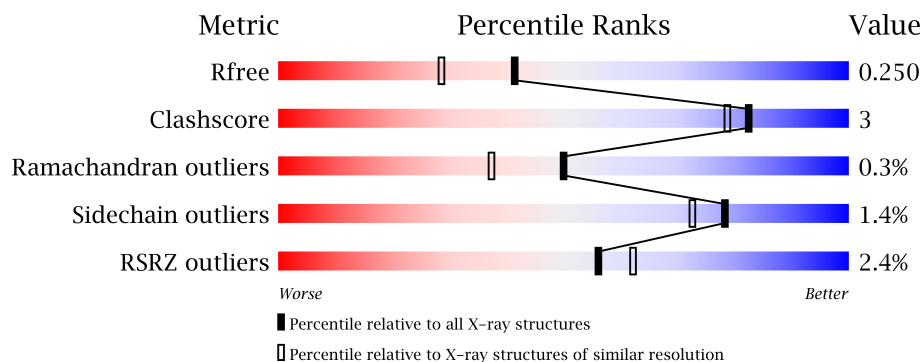
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable24103  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.1.3  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable24103

# 1 Overall quality at a glance

The reported resolution of this entry is 1.94 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	2024 (1.96-1.92)
Clashscore	79885	2281 (1.96-1.92)
Ramachandran outliers	78287	2255 (1.96-1.92)
Sidechain outliers	78261	2255 (1.96-1.92)
RSRZ outliers	66119	2024 (1.96-1.92)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	469	
1	B	469	
1	E	469	
1	G	469	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	K	B	501	-	X
2	K	G	501	-	X
5	1DD	A	506	-	X
5	1DD	B	504	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
5	1DD	E	504	-	X
5	1DD	G	505	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 15234 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Aldehyde dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	446	Total	C	N	O	S	0	6	0
			3519	2242	598	661	18			
1	B	447	Total	C	N	O	S	0	2	0
			3498	2230	593	657	18			
1	E	447	Total	C	N	O	S	0	6	0
			3521	2244	598	661	18			
1	G	447	Total	C	N	O	S	0	5	0
			3509	2236	594	661	18			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-16	HIS	-	EXPRESSION TAG	UNP P30838
A	-15	HIS	-	EXPRESSION TAG	UNP P30838
A	-14	HIS	-	EXPRESSION TAG	UNP P30838
A	-13	HIS	-	EXPRESSION TAG	UNP P30838
A	-12	HIS	-	EXPRESSION TAG	UNP P30838
A	-11	HIS	-	EXPRESSION TAG	UNP P30838
A	-10	SER	-	EXPRESSION TAG	UNP P30838
A	-9	SER	-	EXPRESSION TAG	UNP P30838
A	-8	GLY	-	EXPRESSION TAG	UNP P30838
A	-7	LEU	-	EXPRESSION TAG	UNP P30838
A	-6	VAL	-	EXPRESSION TAG	UNP P30838
A	-5	PRO	-	EXPRESSION TAG	UNP P30838
A	-4	ARG	-	EXPRESSION TAG	UNP P30838
A	-3	GLY	-	EXPRESSION TAG	UNP P30838
A	-2	SER	-	EXPRESSION TAG	UNP P30838
A	-1	HIS	-	EXPRESSION TAG	UNP P30838
A	133	ALA	SER	NATURAL VARIANT	UNP P30838
B	-16	HIS	-	EXPRESSION TAG	UNP P30838
B	-15	HIS	-	EXPRESSION TAG	UNP P30838
B	-14	HIS	-	EXPRESSION TAG	UNP P30838
B	-13	HIS	-	EXPRESSION TAG	UNP P30838

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-12	HIS	-	EXPRESSION TAG	UNP P30838
B	-11	HIS	-	EXPRESSION TAG	UNP P30838
B	-10	SER	-	EXPRESSION TAG	UNP P30838
B	-9	SER	-	EXPRESSION TAG	UNP P30838
B	-8	GLY	-	EXPRESSION TAG	UNP P30838
B	-7	LEU	-	EXPRESSION TAG	UNP P30838
B	-6	VAL	-	EXPRESSION TAG	UNP P30838
B	-5	PRO	-	EXPRESSION TAG	UNP P30838
B	-4	ARG	-	EXPRESSION TAG	UNP P30838
B	-3	GLY	-	EXPRESSION TAG	UNP P30838
B	-2	SER	-	EXPRESSION TAG	UNP P30838
B	-1	HIS	-	EXPRESSION TAG	UNP P30838
B	133	ALA	SER	NATURAL VARIANT	UNP P30838
E	-16	HIS	-	EXPRESSION TAG	UNP P30838
E	-15	HIS	-	EXPRESSION TAG	UNP P30838
E	-14	HIS	-	EXPRESSION TAG	UNP P30838
E	-13	HIS	-	EXPRESSION TAG	UNP P30838
E	-12	HIS	-	EXPRESSION TAG	UNP P30838
E	-11	HIS	-	EXPRESSION TAG	UNP P30838
E	-10	SER	-	EXPRESSION TAG	UNP P30838
E	-9	SER	-	EXPRESSION TAG	UNP P30838
E	-8	GLY	-	EXPRESSION TAG	UNP P30838
E	-7	LEU	-	EXPRESSION TAG	UNP P30838
E	-6	VAL	-	EXPRESSION TAG	UNP P30838
E	-5	PRO	-	EXPRESSION TAG	UNP P30838
E	-4	ARG	-	EXPRESSION TAG	UNP P30838
E	-3	GLY	-	EXPRESSION TAG	UNP P30838
E	-2	SER	-	EXPRESSION TAG	UNP P30838
E	-1	HIS	-	EXPRESSION TAG	UNP P30838
E	133	ALA	SER	NATURAL VARIANT	UNP P30838
G	-16	HIS	-	EXPRESSION TAG	UNP P30838
G	-15	HIS	-	EXPRESSION TAG	UNP P30838
G	-14	HIS	-	EXPRESSION TAG	UNP P30838
G	-13	HIS	-	EXPRESSION TAG	UNP P30838
G	-12	HIS	-	EXPRESSION TAG	UNP P30838
G	-11	HIS	-	EXPRESSION TAG	UNP P30838
G	-10	SER	-	EXPRESSION TAG	UNP P30838
G	-9	SER	-	EXPRESSION TAG	UNP P30838
G	-8	GLY	-	EXPRESSION TAG	UNP P30838
G	-7	LEU	-	EXPRESSION TAG	UNP P30838
G	-6	VAL	-	EXPRESSION TAG	UNP P30838
G	-5	PRO	-	EXPRESSION TAG	UNP P30838

*Continued on next page...*

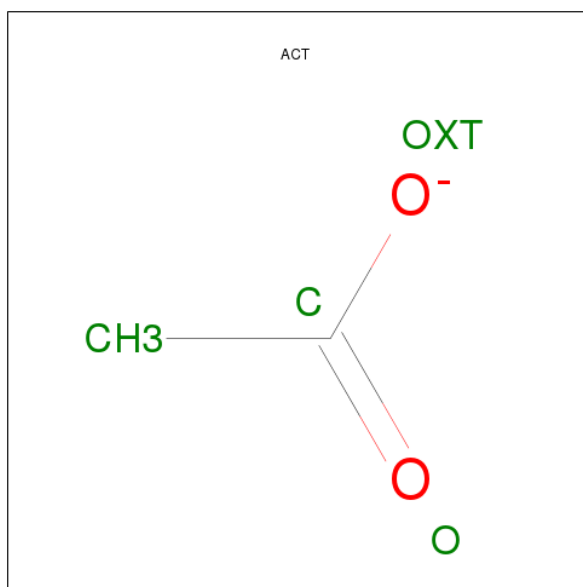
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	-4	ARG	-	EXPRESSION TAG	UNP P30838
G	-3	GLY	-	EXPRESSION TAG	UNP P30838
G	-2	SER	-	EXPRESSION TAG	UNP P30838
G	-1	HIS	-	EXPRESSION TAG	UNP P30838
G	133	ALA	SER	NATURAL VARIANT	UNP P30838

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	2	Total K 2 2	0	0
2	B	2	Total K 2 2	0	0
2	A	2	Total K 2 2	0	0
2	E	1	Total K 1 1	0	0

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



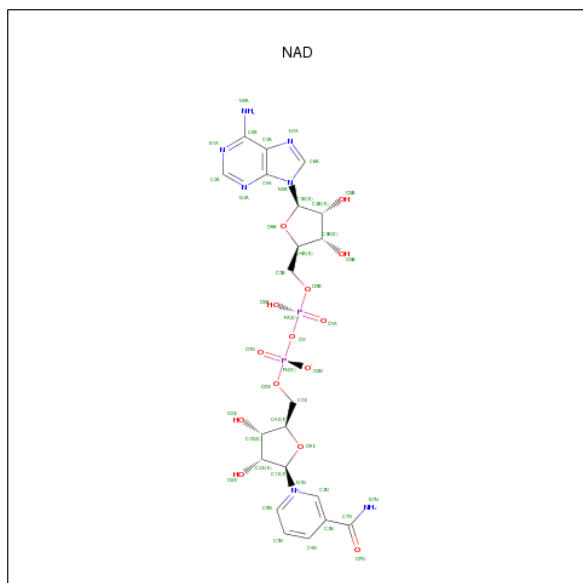
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0

Continued on next page...

Continued from previous page...

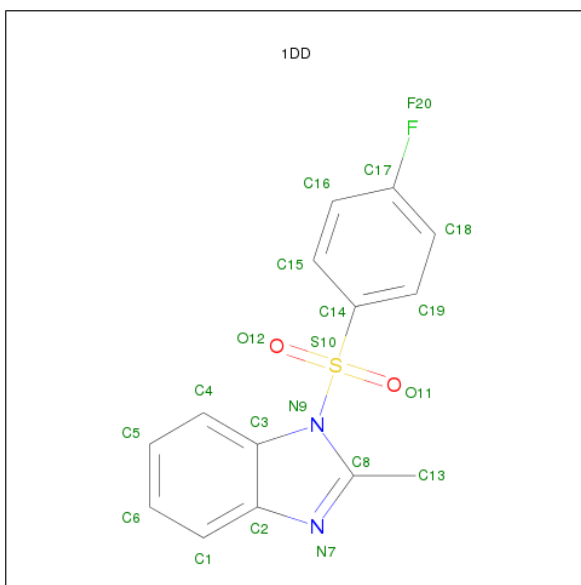
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total	C	O	0	0
			4	2	2		
3	G	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula:  $C_{21}H_{27}N_7O_{14}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	B	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	E	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	G	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

- Molecule 5 is 1-[(4-FLUOROPHENYL)SULFONYL]-2-METHYL-1H-BENZIMIDAZOLE (three-letter code: 1DD) (formula:  $C_{14}H_{11}FN_2O_2S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	F	N	O	S	
			20	14	1	2	2	1	0
5	B	1	Total	C	F	N	O	S	
			20	14	1	2	2	1	0
5	E	1	Total	C	F	N	O	S	
			20	14	1	2	2	1	0
5	G	1	Total	C	F	N	O	S	
			20	14	1	2	2	1	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	212	Total	O		
			212	212	0	0
6	B	256	Total	O		
			256	256	0	0
6	E	235	Total	O		
			235	235	0	0
6	G	205	Total	O		
			205	205	0	0

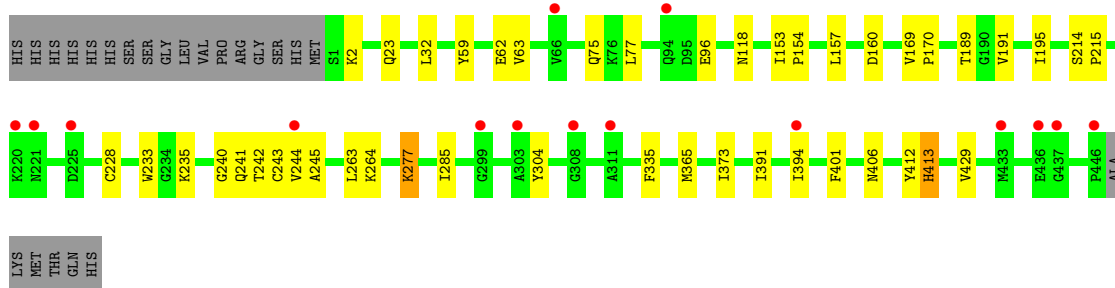


### 3 Residue-property plots

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.

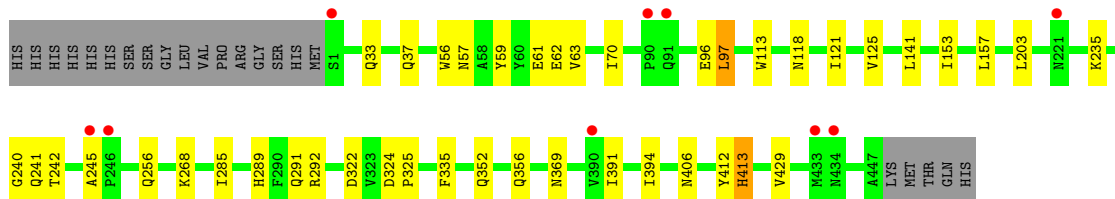
- Molecule 1: Aldehyde dehydrogenase

Chain A: 



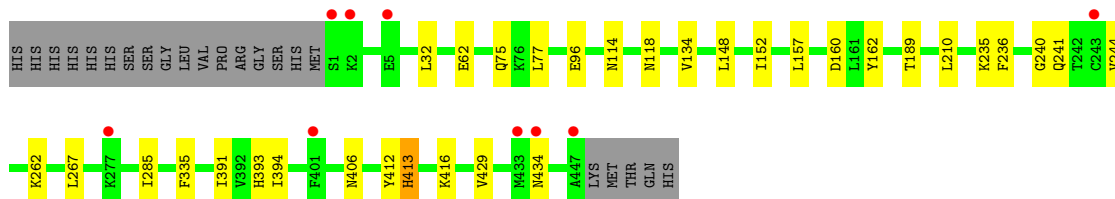
- Molecule 1: Aldehyde dehydrogenase

Chain B: 



- Molecule 1: Aldehyde dehydrogenase

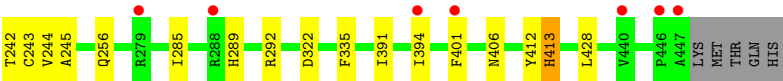
Chain E: 



- Molecule 1: Aldehyde dehydrogenase

Chain G: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	95.22Å 90.92Å 117.92Å 90.00° 112.48° 90.00°	Depositor
Resolution (Å)	40.00 – 1.94 41.95 – 1.94	Depositor EDS
% Data completeness (in resolution range)	95.7 (40.00-1.94) 95.7 (41.95-1.94)	Depositor EDS
$R_{merge}$	0.34	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.63 (at 1.94Å)	Xtriage
Refinement program	REFMAC 5.0	Depositor
R, $R_{free}$	0.215 , 0.250 0.215 , 0.250	Depositor DCC
$R_{free}$ test set	6642 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.1	Xtriage
Anisotropy	0.178	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 28.4	EDS
Estimated twinning fraction	0.011 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	2 of 132117 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	15234	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.66 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 5.4431e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 1DD, K, NAD, ACT

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	0.49	1/3604 (0.0%)	0.49	1/4885 (0.0%)
1	B	0.39	1/3577 (0.0%)	0.48	1/4850 (0.0%)
1	E	0.39	0/3609	0.48	1/4892 (0.0%)
1	G	0.39	0/3594	0.48	0/4873
All	All	0.41	2/14384 (0.0%)	0.48	3/19500 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	56	TRP	CD2-CE2	5.15	1.47	1.41
1	A	233	TRP	CD2-CE2	5.04	1.47	1.41

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	97	LEU	CA-CB-CG	5.44	127.81	115.30
1	E	160	ASP	CB-CG-OD2	5.22	123.00	118.30
1	A	160	ASP	CB-CG-OD2	5.16	122.94	118.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3519	0	3559	25	0
1	B	3498	0	3535	21	0
1	E	3521	0	3565	17	0
1	G	3509	0	3547	26	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	E	1	0	0	0	0
2	G	2	0	0	0	0
3	A	8	0	6	0	0
3	E	4	0	3	0	0
3	G	4	0	3	0	0
4	A	44	0	26	0	0
4	B	44	0	26	1	0
4	E	44	0	26	0	0
4	G	44	0	26	1	0
5	A	20	0	11	3	0
5	B	20	0	11	1	0
5	E	20	0	11	1	0
5	G	20	0	11	4	0
6	A	212	0	0	1	0
6	B	256	0	0	0	0
6	E	235	0	0	2	0
6	G	205	0	0	1	0
All	All	15234	0	14366	90	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

All (90) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:G:256:GLN:HE22	1:G:322:ASP:H	1.19	0.91
1:B:256:GLN:HE22	1:B:322:ASP:H	1.26	0.83
1:B:289:HIS:HD2	1:B:292:ARG:HE	1.42	0.67
1:G:412:TYR:O	1:G:413:HIS:HB2	1.94	0.67
1:G:289:HIS:HD2	1:G:292:ARG:HE	1.44	0.64
1:G:53:LYS:HE3	1:G:238:ASN:ND2	2.12	0.64
1:E:412:TYR:O	1:E:413:HIS:HB2	1.96	0.63
1:A:412:TYR:O	1:A:413:HIS:HB2	1.99	0.63
1:A:241:GLN:HE22	1:A:285:ILE:H	1.48	0.61
1:A:243:CYS:SG	6:A:812:HOH:O	2.15	0.60
1:G:53:LYS:HE3	1:G:238:ASN:HD22	1.67	0.59
1:G:241:GLN:HE22	1:G:285:ILE:H	1.47	0.59
1:B:412:TYR:O	1:B:413:HIS:HB2	2.01	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:62:GLU:HG3	1:B:118:ASN:HB2	1.84	0.58
1:A:62:GLU:HG3	1:A:118:ASN:HB2	1.85	0.58
1:B:241:GLN:HE22	1:B:285:ILE:H	1.50	0.57
1:B:33:GLN:HG3	1:B:70:ILE:HG21	1.86	0.57
1:E:391:ILE:O	1:E:394:ILE:HG12	2.05	0.56
1:A:391:ILE:O	1:A:394:ILE:HG12	2.05	0.56
1:G:244:VAL:HG22	6:G:779:HOH:O	2.05	0.56
1:A:264:LYS:HE2	1:A:304:TYR:CD1	2.41	0.56
1:G:256:GLN:NE2	1:G:322:ASP:H	1.97	0.55
1:B:256:GLN:NE2	1:B:322:ASP:H	2.00	0.54
1:A:96:GLU:HB2	1:A:429:VAL:HB	1.89	0.54
1:B:391:ILE:O	1:B:394:ILE:HG12	2.08	0.54
1:B:352:GLN:O	1:B:356:GLN:HG2	2.07	0.53
1:G:401:PHE:CE2	5:G:505:1DD:H11	2.44	0.53
1:G:243:CYS:SG	4:G:504:NAD:C4N	2.97	0.52
1:G:391:ILE:O	1:G:394:ILE:HG12	2.10	0.52
1:E:32[A]:LEU:HG	1:E:157:LEU:HD21	1.92	0.52
1:E:241:GLN:HE22	1:E:285:ILE:H	1.56	0.52
1:E:96:GLU:HB2	1:E:429:VAL:HB	1.92	0.52
1:G:244:VAL:HG23	1:G:391:ILE:HD13	1.92	0.52
1:G:32:LEU:HG	1:G:157:LEU:HD21	1.93	0.51
1:A:32[A]:LEU:HG	1:A:157:LEU:HD21	1.92	0.50
1:A:242:THR:HB	1:A:245:ALA:HB2	1.94	0.50
1:B:96:GLU:HB2	1:B:429:VAL:HB	1.95	0.48
1:A:394:ILE:HD12	5:A:506:1DD:C2	2.43	0.48
1:E:394:ILE:HD12	5:E:504:1DD:C2	2.44	0.47
1:A:401:PHE:CE2	5:A:506:1DD:H11	2.49	0.47
1:E:244:VAL:HG21	1:E:394:ILE:HD13	1.96	0.47
1:G:394:ILE:HD12	5:G:505:1DD:C2	2.45	0.47
1:B:394:ILE:HD12	5:B:504:1DD:C2	2.44	0.47
1:G:235:LYS:O	1:G:240:GLY:HA2	2.15	0.46
1:A:153:ILE:HB	1:A:154:PRO:HD3	1.97	0.46
1:E:393:HIS:HE1	6:E:771:HOH:O	1.98	0.46
1:A:235:LYS:O	1:A:240:GLY:HA2	2.16	0.46
1:A:191:VAL:O	1:A:195:ILE:HG12	2.16	0.46
1:A:153:ILE:HG23	1:A:157:LEU:HD12	1.98	0.45
1:A:412:TYR:O	1:A:413:HIS:CB	2.65	0.45
1:G:97:LEU:CD2	1:G:428:LEU:HD13	2.47	0.45
1:G:412:TYR:O	1:G:413:HIS:CB	2.64	0.45
1:G:121:ILE:O	1:G:125:VAL:HG23	2.17	0.44
1:B:235:LYS:O	1:B:240:GLY:HA2	2.18	0.44
1:A:59:TYR:HA	1:A:63:VAL:HB	1.99	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:G:33:GLN:HG3	1:G:70:ILE:HG21	2.00	0.43
1:E:235:LYS:O	1:E:240:GLY:HA2	2.19	0.43
1:G:289:HIS:CD2	1:G:292:ARG:HE	2.31	0.43
1:G:89:THR:HB	1:G:90:PRO:HD2	2.01	0.42
1:A:277:LYS:N	1:A:277:LYS:HD2	2.33	0.42
1:E:236:PHE:CD2	1:E:267:LEU:HD21	2.53	0.42
1:B:153:ILE:HG23	1:B:157:LEU:HD12	2.01	0.42
1:A:228:CYS:HB3	1:A:263:LEU:HD23	2.00	0.42
1:A:23[A]:GLN:HA	1:A:23[A]:GLN:OE1	2.18	0.42
1:E:114:ASN:ND2	6:E:808:HOH:O	2.47	0.42
1:E:148:LEU:O	1:E:152:ILE:HG12	2.19	0.42
1:B:113:TRP:CD1	1:B:141:LEU:HD12	2.54	0.42
1:A:244:VAL:HG21	1:A:394:ILE:HD13	2.02	0.42
1:G:62:GLU:HG3	1:G:118:ASN:HB2	2.01	0.42
1:A:365:MET:SD	1:A:373:ILE:HG23	2.59	0.41
1:G:153:ILE:HB	1:G:154:PRO:HD3	2.01	0.41
5:G:505:1DD:C13	5:G:505:1DD:O12	2.68	0.41
1:G:242:THR:HB	1:G:245:ALA:HB2	2.02	0.41
1:A:189:THR:HG23	1:B:203:LEU:HD11	2.03	0.41
1:G:120:THR:HB	1:G:136[A]:LEU:HD11	2.03	0.41
1:E:134:VAL:O	1:E:162:TYR:HA	2.19	0.41
5:A:506:1DD:C13	5:A:506:1DD:O12	2.68	0.41
1:B:324:ASP:HA	1:B:325:PRO:HD2	1.96	0.41
1:B:57:ASN:O	1:B:61:GLU:HB2	2.20	0.41
1:E:62:GLU:HG3	1:E:118:ASN:HB2	2.03	0.41
1:B:121:ILE:O	1:B:125:VAL:HG23	2.20	0.41
1:A:214:SER:HA	1:A:215:PRO:HD3	1.94	0.41
1:E:412:TYR:O	1:E:413:HIS:CB	2.64	0.41
1:G:65:TYR:CD2	5:G:505:1DD:H4	2.56	0.41
1:B:242:THR:HB	1:B:245:ALA:HB2	2.02	0.40
1:B:59:TYR:HA	1:B:63:VAL:HB	2.03	0.40
1:A:169:VAL:HB	1:A:170:PRO:HD3	2.03	0.40
1:E:416:LYS:HD2	1:E:416:LYS:HA	1.89	0.40
1:B:289:HIS:HE1	4:B:503:NAD:O2A	2.05	0.40
1:E:189:THR:HA	1:E:210:LEU:HD13	2.03	0.40

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	
1	A	450/469 (96%)	440 (98%)	9 (2%)	1 (0%)	56	45
1	B	447/469 (95%)	436 (98%)	10 (2%)	1 (0%)	56	45
1	E	451/469 (96%)	444 (98%)	5 (1%)	2 (0%)	43	29
1	G	450/469 (96%)	444 (99%)	5 (1%)	1 (0%)	56	45
All	All	1798/1876 (96%)	1764 (98%)	29 (2%)	5 (0%)	50	37

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	413	HIS
1	B	413	HIS
1	E	413	HIS
1	G	413	HIS
1	E	434	ASN

### 5.3.2 Protein sidechains ⓘ

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	386/400 (96%)	380 (98%)	6 (2%)	75	68
1	B	382/400 (96%)	375 (98%)	7 (2%)	71	63
1	E	386/400 (96%)	381 (99%)	5 (1%)	80	75
1	G	385/400 (96%)	381 (99%)	4 (1%)	85	83
All	All	1539/1600 (96%)	1517 (99%)	22 (1%)	78	72



All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	LYS
1	A	75	GLN
1	A	77	LEU
1	A	277	LYS
1	A	335	PHE
1	A	406	ASN
1	B	37	GLN
1	B	97	LEU
1	B	268	LYS
1	B	291	GLN
1	B	335	PHE
1	B	369	ASN
1	B	406	ASN
1	E	75	GLN
1	E	77	LEU
1	E	262	LYS
1	E	335	PHE
1	E	406	ASN
1	G	37	GLN
1	G	75	GLN
1	G	335	PHE
1	G	406	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (39) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	27	GLN
1	A	118	ASN
1	A	241	GLN
1	A	291	GLN
1	A	393	HIS
1	A	397	HIS
1	A	423	HIS
1	B	33	GLN
1	B	37	GLN
1	B	93	GLN
1	B	118	ASN
1	B	122	GLN
1	B	238	ASN
1	B	241	GLN
1	B	256	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	289	HIS
1	B	291	GLN
1	B	369	ASN
1	B	406	ASN
1	B	423	HIS
1	E	27	GLN
1	E	37	GLN
1	E	241	GLN
1	E	393	HIS
1	E	397	HIS
1	E	406	ASN
1	E	423	HIS
1	G	27	GLN
1	G	33	GLN
1	G	37	GLN
1	G	75	GLN
1	G	118	ASN
1	G	122	GLN
1	G	238	ASN
1	G	241	GLN
1	G	256	GLN
1	G	289	HIS
1	G	406	ASN
1	G	423	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

### 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 5.6 Ligand geometry ⓘ

Of 19 ligands modelled in this entry, 7 are monoatomic - leaving 12 for Mogul analysis.

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	$\# Z  > 2$
3	ACT	A	503	-	1,3,3	1.08	0	0,3,3	0.00	-
3	ACT	A	504	-	1,3,3	1.14	0	0,3,3	0.00	-
4	NAD	A	505	-	48,48,48	0.93	2 (4%)	73,73,73	1.67	9 (12%)
5	1DD	A	506	-	22,22,22	2.12	3 (13%)	30,33,33	1.80	6 (20%)
4	NAD	B	503	-	48,48,48	0.94	2 (4%)	73,73,73	1.71	9 (12%)
5	1DD	B	504	-	22,22,22	2.11	3 (13%)	30,33,33	1.76	6 (20%)
3	ACT	E	502	-	1,3,3	1.07	0	0,3,3	0.00	-
4	NAD	E	503	-	48,48,48	0.93	2 (4%)	73,73,73	1.70	8 (10%)
5	1DD	E	504	-	22,22,22	2.13	3 (13%)	30,33,33	1.74	5 (16%)
3	ACT	G	503	-	1,3,3	1.11	0	0,3,3	0.00	-
4	NAD	G	504	-	48,48,48	0.94	2 (4%)	73,73,73	1.71	8 (10%)
5	1DD	G	505	-	22,22,22	2.13	3 (13%)	30,33,33	1.76	6 (20%)

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
3	ACT	A	503	-	-	0/0/0/0	0/0/0/0
3	ACT	A	504	-	-	0/0/0/0	0/0/0/0
4	NAD	A	505	-	-	0/30/62/62	0/5/5/5
5	1DD	A	506	-	-	0/6/12/12	0/3/3/3
4	NAD	B	503	-	-	0/30/62/62	0/5/5/5
5	1DD	B	504	-	-	0/6/12/12	0/3/3/3
3	ACT	E	502	-	-	0/0/0/0	0/0/0/0
4	NAD	E	503	-	-	0/30/62/62	0/5/5/5
5	1DD	E	504	-	-	0/6/12/12	0/3/3/3
3	ACT	G	503	-	-	0/0/0/0	0/0/0/0
4	NAD	G	504	-	-	0/30/62/62	0/5/5/5
5	1DD	G	505	-	-	0/6/12/12	0/3/3/3

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	504	1DD	C14-S10	-7.65	1.59	1.77
5	A	506	1DD	C14-S10	-7.62	1.59	1.77
5	G	505	1DD	C14-S10	-7.62	1.59	1.77
5	B	504	1DD	C14-S10	-7.58	1.60	1.77
5	E	504	1DD	C3-C2	3.90	1.48	1.40
5	G	505	1DD	C3-C2	3.86	1.48	1.40
5	A	506	1DD	C3-C2	3.84	1.48	1.40
5	B	504	1DD	C3-C2	3.81	1.48	1.40
5	G	505	1DD	C3-N9	-3.63	1.32	1.39
5	B	504	1DD	C3-N9	-3.63	1.32	1.39
5	E	504	1DD	C3-N9	-3.61	1.32	1.39
5	A	506	1DD	C3-N9	-3.59	1.32	1.39
4	G	504	NAD	C5A-C4A	3.26	1.47	1.40
4	B	503	NAD	C5A-C4A	3.20	1.47	1.40
4	A	505	NAD	C5A-C4A	3.18	1.47	1.40
4	E	503	NAD	C5A-C4A	3.13	1.47	1.40
4	A	505	NAD	C4A-N9A	-2.87	1.33	1.37
4	B	503	NAD	C4A-N9A	-2.85	1.33	1.37
4	E	503	NAD	C4A-N9A	-2.68	1.33	1.37
4	G	504	NAD	C4A-N9A	-2.64	1.33	1.37

All (57) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	504	NAD	C5A-C4A-N3A	-7.34	118.82	125.98
4	E	503	NAD	C5A-C4A-N3A	-7.18	118.98	125.98
4	B	503	NAD	C5A-C4A-N3A	-6.80	119.35	125.98
4	A	505	NAD	N3A-C2A-N1A	-6.72	122.98	128.89
4	A	505	NAD	C5A-C4A-N3A	-6.66	119.48	125.98
4	B	503	NAD	N3A-C2A-N1A	-6.62	123.07	128.89
4	E	503	NAD	N3A-C2A-N1A	-6.55	123.13	128.89
4	G	504	NAD	N3A-C2A-N1A	-6.50	123.17	128.89
4	G	504	NAD	N3A-C4A-N9A	5.64	135.06	125.39
4	E	503	NAD	N3A-C4A-N9A	5.58	134.96	125.39
4	B	503	NAD	N3A-C4A-N9A	5.51	134.84	125.39
4	A	505	NAD	N3A-C4A-N9A	5.46	134.76	125.39
5	A	506	1DD	C14-S10-N9	5.15	112.18	104.35
5	E	504	1DD	C14-S10-N9	4.56	111.28	104.35
5	G	505	1DD	C14-S10-N9	4.50	111.20	104.35
5	B	504	1DD	C8-N9-C3	4.46	114.51	106.08
4	B	503	NAD	O4D-C1D-N1N	4.45	113.02	108.13
5	E	504	1DD	C8-N9-C3	4.36	114.33	106.08
5	A	506	1DD	C8-N9-C3	4.34	114.29	106.08

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	G	505	1DD	C8-N9-C3	4.33	114.27	106.08
4	E	503	NAD	O4D-C1D-N1N	4.29	112.84	108.13
5	B	504	1DD	C14-S10-N9	4.28	110.86	104.35
4	A	505	NAD	O4D-C1D-N1N	4.23	112.78	108.13
4	G	504	NAD	O4D-C1D-N1N	4.02	112.55	108.13
5	E	504	1DD	C4-C3-N9	3.34	136.67	130.60
5	B	504	1DD	C4-C3-N9	3.29	136.58	130.60
5	A	506	1DD	O11-S10-O12	-3.25	109.20	116.68
5	A	506	1DD	C4-C3-N9	3.24	136.49	130.60
5	G	505	1DD	O11-S10-O12	-3.22	109.28	116.68
5	B	504	1DD	O11-S10-O12	-3.21	109.30	116.68
5	G	505	1DD	C4-C3-N9	3.21	136.43	130.60
5	E	504	1DD	O11-S10-O12	-3.15	109.44	116.68
4	E	503	NAD	C4A-C5A-N7A	-2.80	106.70	109.41
4	G	504	NAD	C4A-C5A-N7A	-2.74	106.76	109.41
4	E	503	NAD	C2A-N3A-C4A	2.73	121.14	113.27
4	G	504	NAD	C2A-N3A-C4A	2.69	121.01	113.27
4	B	503	NAD	C8A-N9A-C4A	2.66	109.12	106.96
4	B	503	NAD	C4A-C5A-N7A	-2.61	106.89	109.41
4	B	503	NAD	C2A-N3A-C4A	2.58	120.69	113.27
4	A	505	NAD	C8A-N9A-C4A	2.57	109.05	106.96
4	A	505	NAD	C2A-N3A-C4A	2.54	120.58	113.27
4	A	505	NAD	PN-O3-PA	-2.46	123.26	133.17
4	A	505	NAD	C4A-C5A-N7A	-2.45	107.04	109.41
4	G	504	NAD	PN-O3-PA	-2.40	123.53	133.17
4	E	503	NAD	C8A-N9A-C4A	2.36	108.88	106.96
4	E	503	NAD	PN-O3-PA	-2.18	124.41	133.17
5	B	504	1DD	C16-C17-C18	-2.15	120.04	122.91
4	A	505	NAD	C2A-N1A-C6A	2.15	122.59	118.76
5	G	505	1DD	C16-C17-C18	-2.14	120.06	122.91
5	A	506	1DD	C13-C8-N9	2.14	127.24	124.64
4	G	504	NAD	C3N-C7N-N7N	2.11	120.13	117.78
4	B	503	NAD	PN-O3-PA	-2.07	124.83	133.17
5	G	505	1DD	C13-C8-N9	2.07	127.16	124.64
4	B	503	NAD	C2A-N1A-C6A	2.04	122.39	118.76
5	A	506	1DD	C16-C17-C18	-2.03	120.20	122.91
5	B	504	1DD	C13-C8-N9	2.03	127.11	124.64
5	E	504	1DD	C16-C17-C18	-2.01	120.23	122.91

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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, 95<sup>th</sup> 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(Å²)	Q<0.9
1	A	446/469 (95%)	0.38	15 (3%)	43	48	21, 32, 49, 55	2 (0%)
1	B	447/469 (95%)	0.19	9 (2%)	62	69	20, 30, 41, 57	0
1	E	447/469 (95%)	0.24	9 (2%)	62	69	22, 30, 39, 54	0
1	G	447/469 (95%)	0.24	9 (2%)	62	69	21, 31, 42, 53	1 (0%)
All	All	1787/1876 (95%)	0.26	42 (2%)	56	62	20, 31, 44, 57	3 (0%)

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	447	ALA	6.0
1	A	220	LYS	4.9
1	E	434	ASN	4.5
1	A	244	VAL	3.5
1	E	447	ALA	3.3
1	G	90	PRO	3.2
1	A	221	ASN	3.1
1	A	446	PRO	3.1
1	G	440	VAL	3.0
1	B	221	ASN	3.0
1	G	446	PRO	3.0
1	E	433	MET	3.0
1	B	1	SER	3.0
1	A	308	GLY	2.9
1	E	1	SER	2.8
1	A	94	GLN	2.8
1	B	90	PRO	2.6
1	A	433	MET	2.5
1	E	277	LYS	2.5
1	A	394	ILE	2.5
1	A	437	GLY	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	2	LYS	2.4
1	G	2	LYS	2.3
1	E	401	PHE	2.3
1	A	303	ALA	2.2
1	G	401	PHE	2.2
1	G	394	ILE	2.2
1	B	246	PRO	2.2
1	G	279	ARG	2.2
1	B	245	ALA	2.2
1	B	390	VAL	2.2
1	A	436	GLU	2.2
1	E	243	CYS	2.2
1	B	91	GLN	2.2
1	A	299	GLY	2.2
1	G	288	ARG	2.1
1	E	5	GLU	2.1
1	B	433	MET	2.1
1	A	225	ASP	2.1
1	B	434	ASN	2.0
1	A	311	ALA	2.0
1	A	66	VAL	2.0

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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, 95<sup>th</sup> 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	1DD	E	504	20/20	0.62	13.47	34,35,35,36	20

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	1DD	A	506	20/20	0.57	12.65	34,34,35,35	20
5	1DD	G	505	20/20	0.53	11.06	35,35,36,36	20
5	1DD	B	504	20/20	0.49	10.32	33,33,34,34	20
2	K	B	501	1/1	0.31	5.00	51,51,51,51	0
2	K	G	501	1/1	0.34	3.50	53,53,53,53	0
2	K	B	502	1/1	0.20	1.16	61,61,61,61	0
2	K	G	502	1/1	0.21	1.07	64,64,64,64	0
2	K	E	501	1/1	0.20	0.87	48,48,48,48	0
4	NAD	B	503	44/44	0.13	0.21	27,31,34,35	0
2	K	A	502	1/1	0.14	0.10	62,62,62,62	0
3	ACT	G	503	4/4	0.14	0.04	38,38,38,39	0
4	NAD	A	505	44/44	0.12	0.04	32,35,43,44	0
3	ACT	E	502	4/4	0.19	0.03	37,39,39,39	0
4	NAD	G	504	44/44	0.12	-0.20	28,36,38,39	0
3	ACT	A	503	4/4	0.14	-0.31	46,46,46,47	0
4	NAD	E	503	44/44	0.11	-0.63	30,37,39,40	0
2	K	A	501	1/1	0.11	-0.83	67,67,67,67	0
3	ACT	A	504	4/4	0.12	-0.89	44,45,45,45	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.