wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 07:18 PM GMT

PDB ID	:	4OAY
Title	:	BldD CTD-c-di-GMP complex
Authors	:	Schumacher, M.A.; Tschowri, N.; Buttner, M.; Brennan, R.G.
Deposited on	:	2014-01-06
Resolution	:	1.95 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

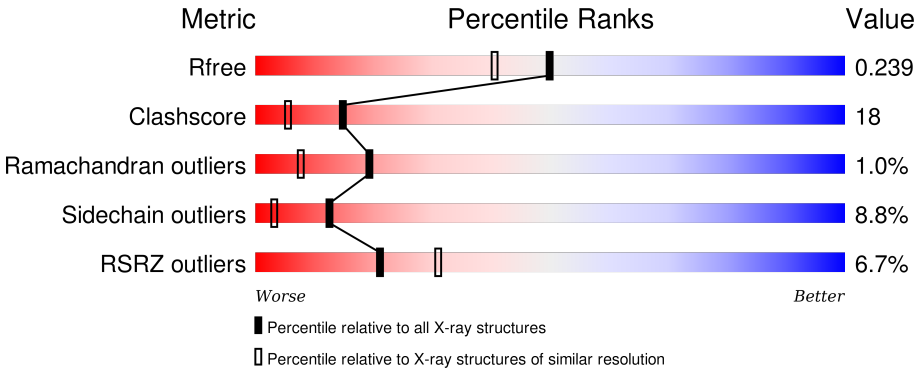
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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.



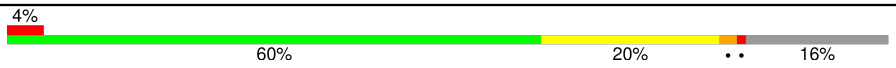
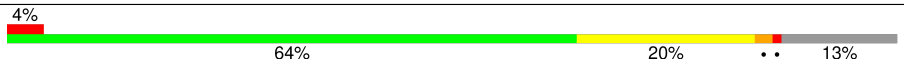
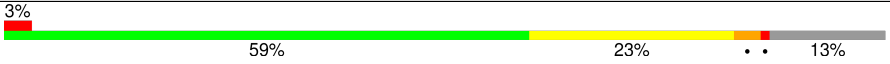
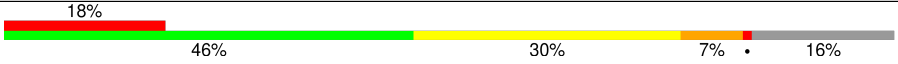
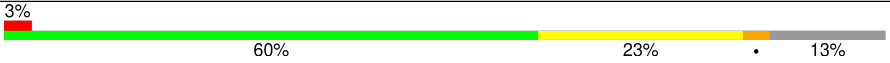
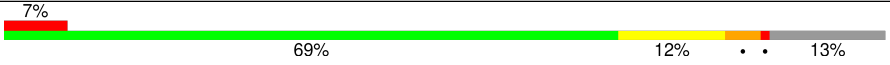
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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

Mol	Chain	Length	Quality of chain
1	A	91	5% 69% 14% • 12%
1	B	91	3% 60% 15% • 20%
1	D	91	7% 71% 14% • 13%
1	E	91	5% 48% 30% • • 16%
1	F	91	2% 59% 20% • 16%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	91	
1	H	91	
1	J	91	
1	K	91	
1	M	91	
1	N	91	
1	R	91	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 16200 atoms, of which 7372 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-binding protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	80	Total	C	H	N	O	Se	0	0	0
			1271	394	642	115	119	1			
1	B	73	Total	C	H	N	O	Se	0	0	0
			1161	362	586	102	110	1			
1	K	79	Total	C	H	N	O	Se	0	0	0
			1257	389	635	114	118	1			
1	J	79	Total	C	H	N	O	Se	0	0	0
			1257	389	635	114	118	1			
1	R	79	Total	C	H	N	O	Se	0	0	0
			1242	386	624	113	118	1			
1	H	76	Total	C	H	N	O	Se	0	0	0
			1192	372	597	107	115	1			
1	E	76	Total	C	H	N	O	Se	0	0	0
			1192	372	597	107	115	1			
1	G	76	Total	C	H	N	O	Se	0	0	0
			1192	372	597	107	115	1			
1	F	76	Total	C	H	N	O	Se	0	0	0
			1207	375	608	108	115	1			
1	D	79	Total	C	H	N	O	Se	0	0	0
			1242	386	624	113	118	1			
1	N	79	Total	C	H	N	O	Se	0	0	0
			1257	389	635	114	118	1			
1	M	76	Total	C	H	N	O	Se	0	0	0
			1190	375	591	108	115	1			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
A	77	SER	-	EXPRESSION TAG	UNP F2RCL8
A	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
A	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
A	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8

Continued on next page...

Continued from previous page...

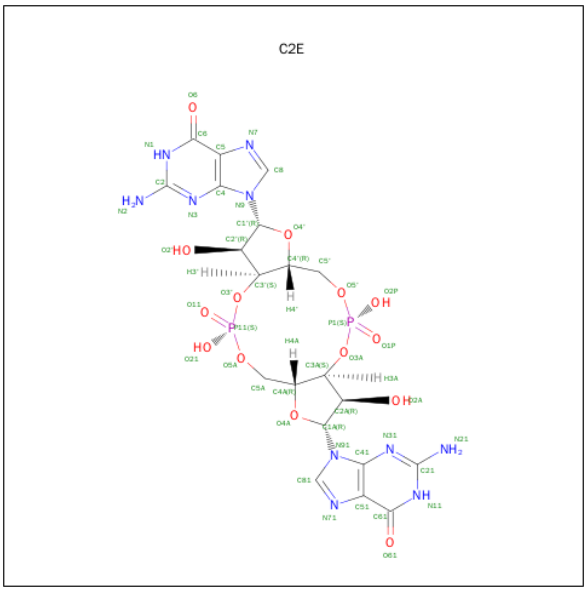
Chain	Residue	Modelled	Actual	Comment	Reference
B	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
B	77	SER	-	EXPRESSION TAG	UNP F2RCL8
B	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
B	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
B	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8
K	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
K	77	SER	-	EXPRESSION TAG	UNP F2RCL8
K	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
K	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
K	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8
J	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
J	77	SER	-	EXPRESSION TAG	UNP F2RCL8
J	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
J	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
J	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8
R	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
R	77	SER	-	EXPRESSION TAG	UNP F2RCL8
R	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
R	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
R	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8
H	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
H	77	SER	-	EXPRESSION TAG	UNP F2RCL8
H	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
H	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
H	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8
E	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
E	77	SER	-	EXPRESSION TAG	UNP F2RCL8
E	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
E	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
E	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8
G	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
G	77	SER	-	EXPRESSION TAG	UNP F2RCL8
G	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
G	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
G	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8
F	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
F	77	SER	-	EXPRESSION TAG	UNP F2RCL8
F	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
F	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
F	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8
D	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
D	77	SER	-	EXPRESSION TAG	UNP F2RCL8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
D	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
D	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8
N	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
N	77	SER	-	EXPRESSION TAG	UNP F2RCL8
N	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
N	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
N	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8
M	76	GLY	-	EXPRESSION TAG	UNP F2RCL8
M	77	SER	-	EXPRESSION TAG	UNP F2RCL8
M	78	HIS	-	EXPRESSION TAG	UNP F2RCL8
M	79	MSE	-	EXPRESSION TAG	UNP F2RCL8
M	92	MSE	LEU	ENGINEERED MUTATION	UNP F2RCL8

- Molecule 2 is 9,9'-[(2R,3R,3AS,5S,7AR,9R,10R,10AS,12S,14AR)-3,5,10,12-TETRAHYDROXY-5,12-DIOXIDOOCTAHYDRO-2H,7H-DIFURO[3,2-D:3',2'-J][1,3,7,9,2,8]TETRAOXADIPHOSPHACYCLODODECINE-2,9-DIYL]BIS(2-AMINO-1,9-DIHYDRO-6H-PURIN-6-ONE) (three-letter code: C2E) (formula: $C_{20}H_{24}N_{10}O_{14}P_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C H N O P 46 19 1 10 14 2	0	0
2	A	1	Total C N O P 46 20 10 14 2	0	0
2	B	1	Total C N O P 46 20 10 14 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	K	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	K	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	J	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	J	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	R	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	R	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	H	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	H	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	E	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	E	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	G	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	G	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	F	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	F	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	D	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	D	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	N	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	N	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	M	1	Total	C	N	O	P	0	0
			46	20	10	14	2		
2	M	1	Total	C	N	O	P	0	0
			46	20	10	14	2		

- Molecule 3 is water.

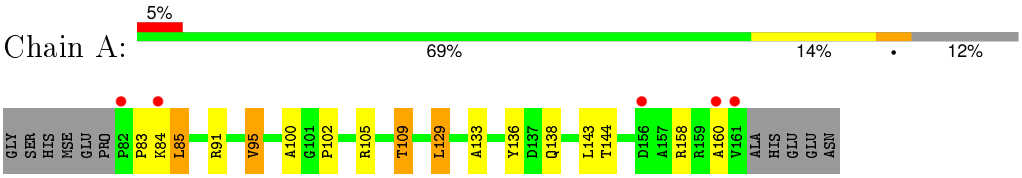
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	23	Total O 23 23	0	0
3	B	32	Total O 32 32	0	0
3	K	53	Total O 53 53	0	0
3	J	50	Total O 50 50	0	0
3	R	58	Total O 58 58	0	0
3	H	20	Total O 20 20	0	0
3	E	26	Total O 26 26	0	0
3	G	56	Total O 56 56	0	0
3	F	27	Total O 27 27	0	0
3	D	46	Total O 46 46	0	0
3	N	34	Total O 34 34	0	0
3	M	11	Total O 11 11	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.

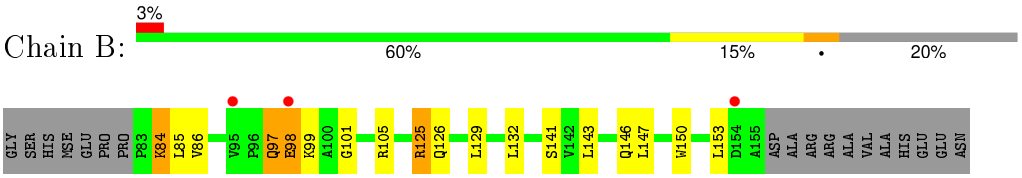
- Molecule 1: DNA-binding protein

Chain A:



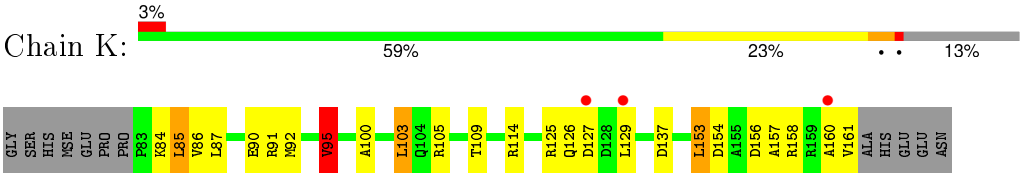
- Molecule 1: DNA-binding protein

Chain B:



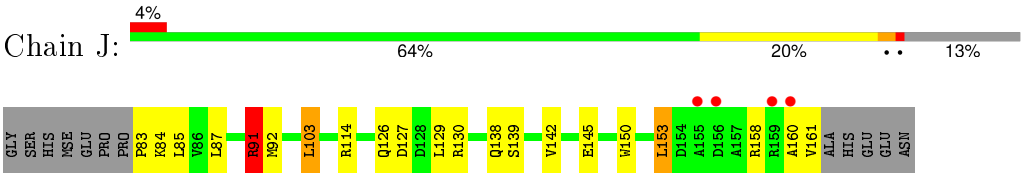
- Molecule 1: DNA-binding protein

Chain K:



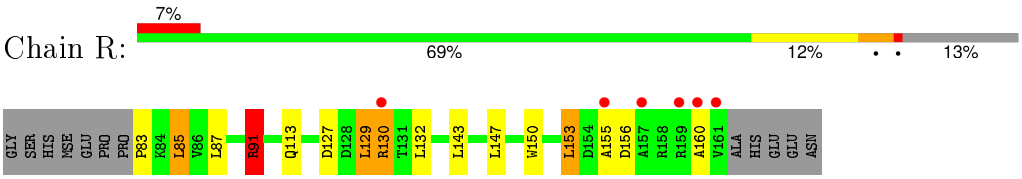
- Molecule 1: DNA-binding protein

Chain J:



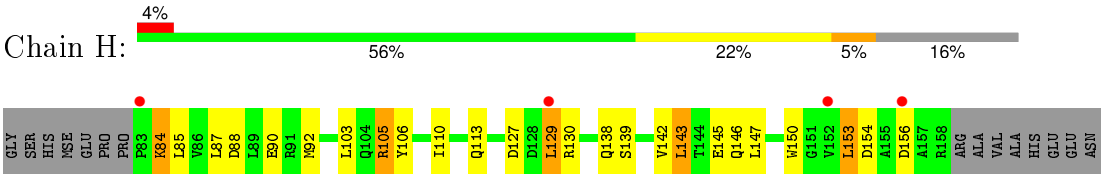
- Molecule 1: DNA-binding protein

Chain R:



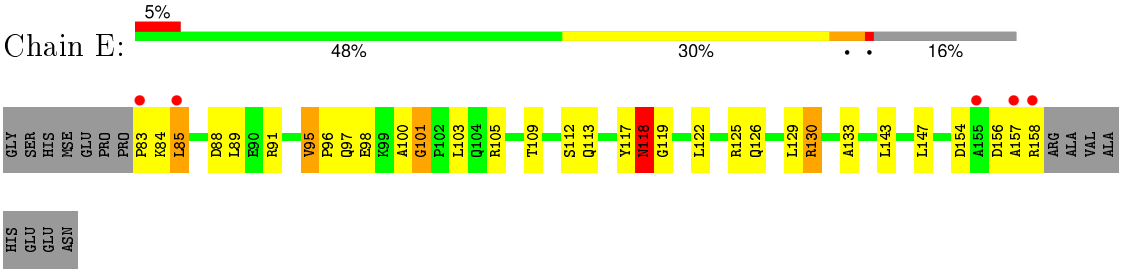
- Molecule 1: DNA-binding protein

Chain H:



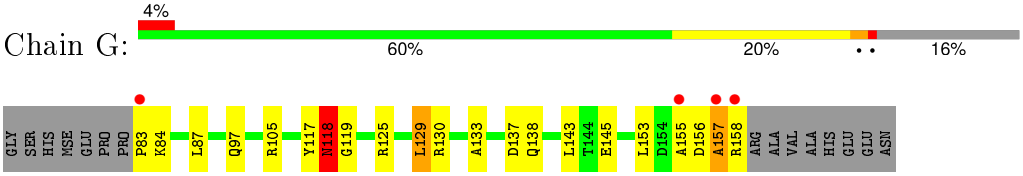
• Molecule 1: DNA-binding protein

Chain E:



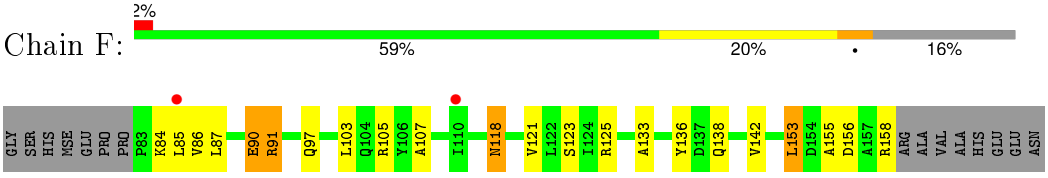
• Molecule 1: DNA-binding protein

Chain G:



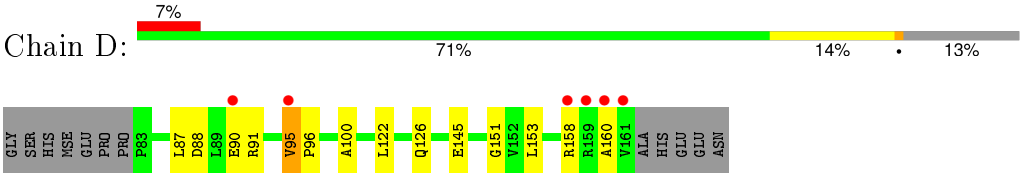
• Molecule 1: DNA-binding protein

Chain F:



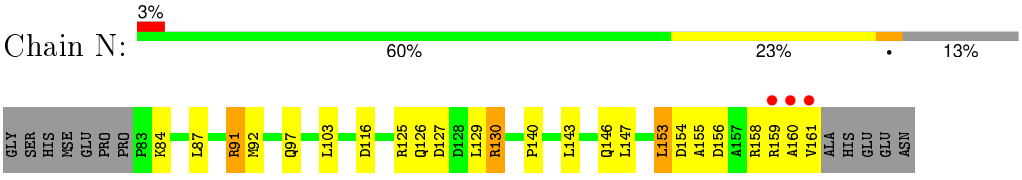
• Molecule 1: DNA-binding protein

Chain D:



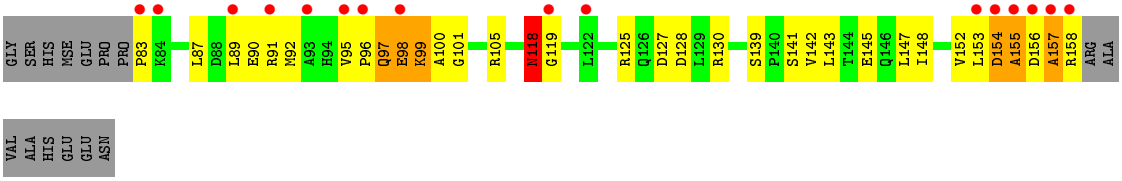
• Molecule 1: DNA-binding protein

Chain N:



• Molecule 1: DNA-binding protein

Chain M:


GLY
SER
HIS
MSE
GLU
PRO
PRO
P83
K84
L87
D88
L89
E90
R91
N92
A93
H94
V95
P96
Q97
E98
K99
A100
G101
R105
N118
G119
L122
R125
Q126
D127
D128
L129
R130
S139
P140
S141
V142
L143
T144
E145
Q146
L147
I148
V152
L153
D154
A155
D156
A157
R158
ARG
ALA
VAL
ALA
HIS
GLU
GLU
ASN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	86.50Å 86.50Å 151.48Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	74.91 – 1.95 74.91 – 1.95	Depositor EDS
% Data completeness (in resolution range)	98.5 (74.91-1.95) 98.5 (74.91-1.95)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	1.60 (at 1.95Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.6.4_486)	Depositor
R, R_{free}	0.198 , 0.234 0.210 , 0.239	Depositor DCC
R_{free} test set	2009 reflections (2.26%)	DCC
Wilson B-factor (Å^2)	39.8	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 46.1	EDS
Estimated twinning fraction	0.006 for -h,-k,l 0.033 for h,-h-k,-l 0.018 for -k,-h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 90989 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	16200	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.21% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: C2E

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	0.71	0/638	0.88	0/864
1	B	0.79	0/583	0.87	1/789 (0.1%)
1	D	0.92	0/626	0.86	0/848
1	E	0.80	0/603	0.87	2/817 (0.2%)
1	F	0.82	1/607 (0.2%)	0.92	2/821 (0.2%)
1	G	1.02	1/603 (0.2%)	0.96	2/817 (0.2%)
1	H	0.81	0/603	0.80	1/817 (0.1%)
1	J	0.96	0/630	0.94	1/852 (0.1%)
1	K	0.94	2/630 (0.3%)	0.95	3/852 (0.4%)
1	M	0.75	0/607	0.81	0/821
1	N	0.95	0/630	0.95	1/852 (0.1%)
1	R	0.97	0/626	0.91	1/848 (0.1%)
All	All	0.88	4/7386 (0.1%)	0.90	14/9998 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	107	ALA	CA-CB	6.39	1.65	1.52
1	G	145	GLU	CG-CD	5.93	1.60	1.51
1	K	90	GLU	CB-CG	5.33	1.62	1.52
1	K	90	GLU	CG-CD	5.08	1.59	1.51

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	N	116	ASP	CB-CG-OD1	7.87	125.38	118.30
1	F	91	ARG	NE-CZ-NH2	-7.51	116.54	120.30
1	F	91	ARG	NE-CZ-NH1	6.84	123.72	120.30
1	B	125	ARG	NE-CZ-NH2	-6.70	116.95	120.30
1	J	91	ARG	NE-CZ-NH1	6.66	123.63	120.30

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	629	642	641	14	0
1	B	575	586	585	13	0
1	D	618	624	623	12	0
1	E	595	597	596	35	1
1	F	599	608	607	21	0
1	G	595	597	596	20	0
1	H	595	597	596	31	1
1	J	622	635	634	24	1
1	K	622	635	634	21	0
1	M	599	591	607	41	0
1	N	622	635	634	22	0
1	R	618	624	623	17	0
2	A	91	1	38	5	0
2	B	92	0	39	3	0
2	D	92	0	41	3	0
2	E	92	0	41	1	0
2	F	92	0	44	2	0
2	G	92	0	41	3	0
2	H	92	0	41	1	0
2	J	92	0	41	0	0
2	K	92	0	41	4	0
2	M	92	0	39	4	0
2	N	92	0	40	3	0
2	R	92	0	43	0	0
3	A	23	0	0	1	0
3	B	32	0	0	1	2
3	D	46	0	0	2	0
3	E	26	0	0	9	0
3	F	27	0	0	4	2
3	G	56	0	0	8	0
3	H	20	0	0	5	0
3	J	50	0	0	2	0
3	K	53	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	M	11	0	0	2	0
3	N	34	0	0	7	0
3	R	58	0	0	4	1
All	All	8828	7372	7865	288	4

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 18.

The worst 5 of 288 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:K:85:LEU:HD13	1:K:129:LEU:HD11	1.29	1.13
1:G:83:PRO:N	3:G:353:HOH:O	1.82	1.10
1:G:155:ALA:HB2	3:G:334:HOH:O	1.61	1.00
1:D:90:GLU:HB3	3:D:343:HOH:O	1.62	1.00
1:N:92:MSE:HE2	1:N:103:LEU:HD21	1.44	0.95

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:304:HOH:O	3:F:308:HOH:O[1_545]	1.72	0.48
3:B:327:HOH:O	3:F:308:HOH:O[1_545]	1.80	0.40
1:H:113:GLN:HE21	1:E:130:ARG:HH21[1_445]	1.31	0.29
1:J:127:ASP:OD1	3:R:325:HOH:O[1_455]	1.92	0.28

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	78/91 (86%)	77 (99%)	1 (1%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	71/91 (78%)	67 (94%)	3 (4%)	1 (1%)	14 4
1	D	77/91 (85%)	73 (95%)	4 (5%)	0	100 100
1	E	74/91 (81%)	69 (93%)	3 (4%)	2 (3%)	6 1
1	F	74/91 (81%)	73 (99%)	1 (1%)	0	100 100
1	G	74/91 (81%)	70 (95%)	2 (3%)	2 (3%)	6 1
1	H	74/91 (81%)	71 (96%)	3 (4%)	0	100 100
1	J	77/91 (85%)	76 (99%)	1 (1%)	0	100 100
1	K	77/91 (85%)	76 (99%)	1 (1%)	0	100 100
1	M	74/91 (81%)	68 (92%)	2 (3%)	4 (5%)	2 0
1	N	77/91 (85%)	76 (99%)	1 (1%)	0	100 100
1	R	77/91 (85%)	75 (97%)	2 (3%)	0	100 100
All	All	904/1092 (83%)	871 (96%)	24 (3%)	9 (1%)	19 8

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	118	ASN
1	G	157	ALA
1	M	98	GLU
1	M	118	ASN
1	M	157	ALA

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	68/75 (91%)	63 (93%)	5 (7%)	17 5
1	B	63/75 (84%)	59 (94%)	4 (6%)	22 8
1	D	66/75 (88%)	64 (97%)	2 (3%)	48 36
1	E	64/75 (85%)	55 (86%)	9 (14%)	4 1
1	F	65/75 (87%)	59 (91%)	6 (9%)	11 3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	G	64/75 (85%)	59 (92%)	5 (8%)	16 4
1	H	64/75 (85%)	56 (88%)	8 (12%)	6 1
1	J	67/75 (89%)	64 (96%)	3 (4%)	34 18
1	K	67/75 (89%)	62 (92%)	5 (8%)	17 5
1	M	65/75 (87%)	57 (88%)	8 (12%)	6 1
1	N	67/75 (89%)	59 (88%)	8 (12%)	6 1
1	R	66/75 (88%)	60 (91%)	6 (9%)	12 3
All	All	786/900 (87%)	717 (91%)	69 (9%)	12 3

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	85	LEU
1	E	154	ASP
1	M	97	GLN
1	E	95	VAL
1	E	112	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	94	HIS
1	M	146	GLN
1	E	126	GLN
1	K	138	GLN
1	G	97	GLN

5.3.3 RNA

There are no RNA molecules 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 (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

24 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	C2E	A	201	-	39,49,52	2.16	9 (23%)	45,77,82	2.24	11 (24%)
2	C2E	A	202	-	42,52,52	2.20	12 (28%)	50,82,82	2.31	14 (28%)
2	C2E	B	201	-	42,52,52	2.44	15 (35%)	50,82,82	3.10	26 (52%)
2	C2E	B	202	-	42,52,52	2.20	12 (28%)	50,82,82	2.32	15 (30%)
2	C2E	D	201	-	42,52,52	2.20	12 (28%)	50,82,82	2.31	16 (32%)
2	C2E	D	202	-	42,52,52	2.19	9 (21%)	50,82,82	2.50	16 (32%)
2	C2E	E	201	-	42,52,52	2.19	10 (23%)	50,82,82	2.50	16 (32%)
2	C2E	E	202	-	42,52,52	2.20	11 (26%)	50,82,82	2.31	16 (32%)
2	C2E	F	201	-	42,52,52	2.33	10 (23%)	50,82,82	2.58	20 (40%)
2	C2E	F	202	-	42,52,52	2.28	11 (26%)	50,82,82	1.86	11 (22%)
2	C2E	G	201	-	42,52,52	2.64	19 (45%)	50,82,82	2.68	19 (38%)
2	C2E	G	202	-	42,52,52	2.20	11 (26%)	50,82,82	2.31	15 (30%)
2	C2E	H	201	-	42,52,52	2.19	10 (23%)	50,82,82	2.50	16 (32%)
2	C2E	H	202	-	42,52,52	2.20	12 (28%)	50,82,82	2.32	16 (32%)
2	C2E	J	201	-	42,52,52	2.89	18 (42%)	50,82,82	2.76	24 (48%)
2	C2E	J	202	-	42,52,52	2.51	14 (33%)	50,82,82	2.89	21 (42%)
2	C2E	K	201	-	42,52,52	2.19	10 (23%)	50,82,82	2.50	16 (32%)
2	C2E	K	202	-	42,52,52	2.20	11 (26%)	50,82,82	2.31	16 (32%)
2	C2E	M	201	-	42,52,52	2.46	19 (45%)	50,82,82	3.25	21 (42%)
2	C2E	M	202	-	42,52,52	2.20	11 (26%)	50,82,82	2.31	15 (30%)
2	C2E	N	201	-	42,52,52	2.63	18 (42%)	50,82,82	2.88	21 (42%)
2	C2E	N	202	-	42,52,52	2.20	11 (26%)	50,82,82	2.31	16 (32%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	C2E	R	201	-	42,52,52	2.55	16 (38%)	50,82,82	2.95	23 (46%)
2	C2E	R	202	-	42,52,52	2.28	10 (23%)	50,82,82	1.86	11 (22%)

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	C2E	A	201	-	-	0/22/62/62	0/5/6/7
2	C2E	A	202	-	-	0/22/62/62	0/6/7/7
2	C2E	B	201	-	-	0/22/62/62	0/6/7/7
2	C2E	B	202	-	-	0/22/62/62	0/6/7/7
2	C2E	D	201	-	-	0/22/62/62	0/6/7/7
2	C2E	D	202	-	-	0/22/62/62	0/6/7/7
2	C2E	E	201	-	-	0/22/62/62	0/6/7/7
2	C2E	E	202	-	-	0/22/62/62	0/6/7/7
2	C2E	F	201	-	-	0/22/62/62	0/6/7/7
2	C2E	F	202	-	-	0/22/62/62	0/6/7/7
2	C2E	G	201	-	-	0/22/62/62	0/6/7/7
2	C2E	G	202	-	-	0/22/62/62	0/6/7/7
2	C2E	H	201	-	-	0/22/62/62	0/6/7/7
2	C2E	H	202	-	-	0/22/62/62	0/6/7/7
2	C2E	J	201	-	-	0/22/62/62	0/6/7/7
2	C2E	J	202	-	-	0/22/62/62	0/6/7/7
2	C2E	K	201	-	-	0/22/62/62	0/6/7/7
2	C2E	K	202	-	-	0/22/62/62	0/6/7/7
2	C2E	M	201	-	-	0/22/62/62	0/6/7/7
2	C2E	M	202	-	-	0/22/62/62	0/6/7/7
2	C2E	N	201	-	-	0/22/62/62	0/6/7/7
2	C2E	N	202	-	-	0/22/62/62	0/6/7/7
2	C2E	R	201	-	-	0/22/62/62	0/6/7/7
2	C2E	R	202	-	-	0/22/62/62	0/6/7/7

The worst 5 of 301 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	R	202	C2E	C8-N7	-6.35	1.22	1.34
2	F	202	C2E	C8-N7	-6.35	1.22	1.34
2	A	201	C2E	C8-N7	-6.05	1.23	1.35
2	B	202	C2E	C2'-C3'	-5.97	1.39	1.53
2	M	202	C2E	C2'-C3'	-5.96	1.39	1.53

The worst 5 of 411 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	201	C2E	C41-C51-N71	-6.99	103.05	109.48
2	D	202	C2E	C41-C51-N71	-6.90	103.14	109.48
2	E	201	C2E	C41-C51-N71	-6.88	103.15	109.48
2	H	201	C2E	C41-C51-N71	-6.87	103.16	109.48
2	K	201	C2E	C41-C51-N71	-6.85	103.18	109.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

18 monomers are involved in 29 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	201	C2E	4	0
2	A	202	C2E	1	0
2	B	201	C2E	1	0
2	B	202	C2E	2	0
2	D	201	C2E	1	0
2	D	202	C2E	2	0
2	E	202	C2E	1	0
2	F	201	C2E	1	0
2	F	202	C2E	1	0
2	G	201	C2E	2	0
2	G	202	C2E	1	0
2	H	202	C2E	1	0
2	K	201	C2E	3	0
2	K	202	C2E	1	0
2	M	201	C2E	1	0
2	M	202	C2E	3	0
2	N	201	C2E	2	0
2	N	202	C2E	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

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^{th} 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	79/91 (86%)	0.67	5 (6%)	23	33	36, 50, 67, 74	0
1	B	72/91 (79%)	0.69	3 (4%)	40	51	31, 45, 71, 80	0
1	D	78/91 (85%)	0.70	6 (7%)	16	25	26, 43, 69, 72	0
1	E	75/91 (82%)	0.74	5 (6%)	21	31	32, 50, 74, 79	0
1	F	75/91 (82%)	0.56	2 (2%)	58	68	30, 45, 71, 78	0
1	G	75/91 (82%)	0.65	4 (5%)	30	41	27, 37, 65, 73	0
1	H	75/91 (82%)	0.69	4 (5%)	30	41	33, 49, 68, 74	0
1	J	78/91 (85%)	0.60	4 (5%)	32	43	25, 36, 68, 73	0
1	K	78/91 (85%)	0.52	3 (3%)	44	56	28, 40, 64, 67	0
1	M	75/91 (82%)	1.16	16 (21%)	1	1	33, 54, 87, 93	0
1	N	78/91 (85%)	0.55	3 (3%)	44	56	28, 40, 63, 70	0
1	R	78/91 (85%)	0.79	6 (7%)	16	25	24, 36, 71, 75	0
All	All	916/1092 (83%)	0.69	61 (6%)	21	31	24, 44, 72, 93	0

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	160	ALA	8.1
1	J	160	ALA	7.9
1	R	161	VAL	7.6
1	R	160	ALA	6.2
1	M	155	ALA	5.2

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 (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	C2E	N	202	46/46	0.93	0.15	1.07	28,36,42,48	0
2	C2E	R	201	46/46	0.98	0.15	0.73	22,27,30,31	0
2	C2E	D	202	46/46	0.94	0.16	0.68	23,27,32,33	0
2	C2E	J	202	46/46	0.97	0.13	0.51	25,30,34,37	0
2	C2E	N	201	46/46	0.98	0.14	0.41	24,28,31,31	0
2	C2E	H	201	46/46	0.96	0.14	0.20	26,32,37,39	0
2	C2E	J	201	46/46	0.98	0.15	0.19	21,27,30,33	0
2	C2E	B	202	46/46	0.95	0.14	0.17	31,38,47,56	0
2	C2E	D	201	46/46	0.93	0.14	0.16	23,30,38,42	0
2	C2E	K	201	46/46	0.96	0.16	0.15	22,27,30,32	0
2	C2E	M	201	46/46	0.97	0.14	0.01	30,39,45,49	0
2	C2E	E	201	46/46	0.96	0.14	-0.01	27,34,37,39	0
2	C2E	F	201	46/46	0.94	0.14	-0.08	33,37,41,46	0
2	C2E	M	202	46/46	0.91	0.13	-0.09	40,49,55,61	0
2	C2E	F	202	46/46	0.93	0.13	-0.11	37,44,48,51	0
2	C2E	H	202	46/46	0.94	0.14	-0.18	33,40,51,53	0
2	C2E	G	201	46/46	0.98	0.14	-0.22	28,31,35,38	0
2	C2E	B	201	46/46	0.98	0.14	-0.30	25,33,38,40	0
2	C2E	A	201	45/46	0.95	0.13	-0.33	33,37,40,44	0
2	C2E	R	202	46/46	0.96	0.13	-0.40	27,32,41,44	0
2	C2E	K	202	46/46	0.96	0.12	-0.43	27,32,40,45	0
2	C2E	A	202	46/46	0.93	0.11	-0.55	39,46,52,59	0
2	C2E	E	202	46/46	0.94	0.12	-0.69	36,45,51,56	0
2	C2E	G	202	46/46	0.94	0.13	-0.78	31,40,52,58	0

6.5 Other polymers (i)

There are no such residues in this entry.