wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 05:16 PM BST

PDB ID	:	1TVC
Title	:	FAD and NADH binding domain of methane monooxygenase reductase from Methylococcus capsulatus (Bath)
Authors	:	Chatwood, L.L.; Mueller, J.; Gross, J.D.; Wagner, G.; Lippard, S.J.
Deposited on	:	2004-06-29

This is a wwPDB NMR 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/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

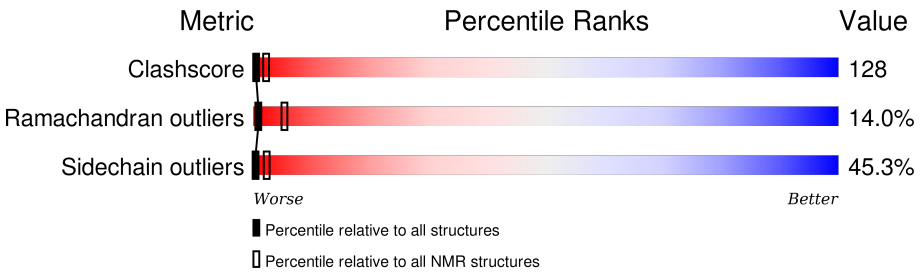
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 75%.

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)	NMR archive (#Entries)
Clashscore	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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%

Mol	Chain	Length	Quality of chain
1	A	250	8% 48% 37% • 6%

2 Ensemble composition and analysis (i)

This entry contains 10 models. Model 4 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:10-A:34, A:43-A:251 (234)	0.83	4

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 1 clusters. No single-model clusters were found.

Cluster number	Models
1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10

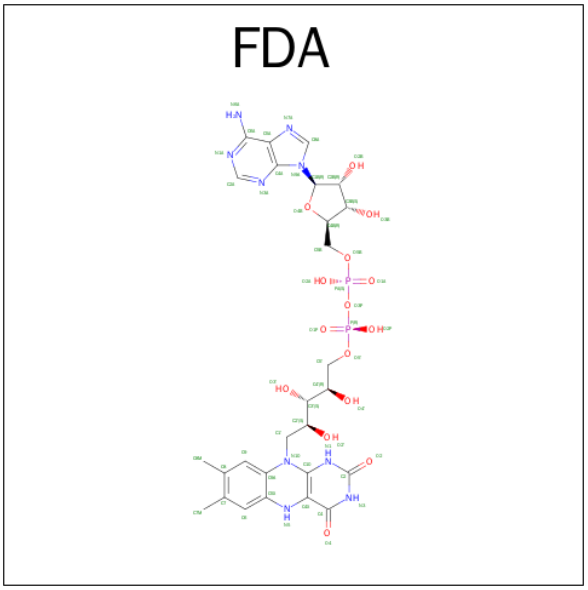
3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3925 atoms, of which 1928 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called METHANE MONOOXYGENASE COMPONENT C.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	250	3839	1229	1895	335	369	11	0

- Molecule 2 is DIHYDROFLAVINE-ADENINE DINUCLEOTIDE (three-letter code: FDA) (formula: $C_{27}H_{35}N_9O_{15}P_2$).



Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	P
2	A	1	86	27	33	9	15	2

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: METHANE MONOOXYGENASE COMPONENT C

Chain A: 8% 48% 37% • 6%

C2 R3 I4 S5 F6 G7 E8 V9 G10 S11 F12 E13 A14 E15 V16 V17 G18 L19 N20 W21 V22 S23 S24 N25 T26 V27 Q28 F29 L30 L31 Q32 K33 R34 P35 D36 E37 C38 G39 N40 R41 G42 V43 K44 F45 E46 P47 G48 Q49 F50 M51 D52 L53 T54 I55 P56 G57 T58 D59 V60 S61

R62 S63 Y64 S65 P66 A67 N68 L69 P70 N71 F72 E73 G74 R75 L76 E77 F78 L79 I80 R81 V82 L83 P84 E85 G86 R87 F88 S89 D90 Y91 L92 R93 N94 D95 A96 R97 V98 G99 Q100 V101 L102 S103 V104 K105 G106 P107 L108 G109 V110 F111 G112 L113 K114 E115 M118 A119 P120 R121 Y122

F123 V124 A125 G126 G127 T128 G129 L130 A131 P132 V133 V134 S135 M136 V137 R138 Q139 M140 Q141 E142 W143 T144 A145 P146 N147 E148 T149 R150 I151 Y152 F153 G154 V155 N156 T157 E158 P159 E160 L161 F162 Y163 I164 D165 E166 L167 K168 S169 L170 E171 R172 S173 M174 R175 N176 L177 T178 V179 K180 A181 C182

V183 W184 H185 P186 S187 G188 D189 W190 E191 G192 E193 Q194 G195 S196 P197 I198 D199 A200 L201 R202 L205 E206 S207 S208 D209 A210 N211 P212 D213 I214 Y215 L216 C217 G218 P219 P220 G221 M222 I223 D224 A225 A226 C227 E228 L229 V230 R231 S232 R233 G234 I235 P236 G237 E238 Q239 V240 F241 F242 E243

K244 F245 L246 P247 S248 G249 A250 A251

4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 4. Colouring as in section 4.1 above.

- Molecule 1: METHANE MONOOXYGENASE COMPONENT C

Chain A: 15% 39% 36% • 6%

C2 R3 I4 S5 F6 G7 E8 V9 F12 E13 V16 V17 G18 L19 N20 W21 V22 S23 S24 N25 T26 V27 Q28 F29 L30 L31 Q32 K33 R34 P35 D36 E37 C38 G39 N40 R41 G42 V43 K44 F45 E46 P47 G48 Q49 F50 M51 D52 L53 T54 I55 P56 G57 T58 D59 V60 S61 R62 S63

Y64 S65 P66 A67 N68 L69 P70 N71 F72 E73 G74 R75 L76 E77 F78 L79 I80 R81 V82 L83 P84 E85 F88 S89 D90 Y91 L92 R93 N94 D95 A96 R97 V98 G99 Q100 V101 L102 S103 V104 K105 L108 G109 V110 F111 G112 L113 K114 E115 M118 R121 Y122 F123 V124 A125 G126 G127

T128 G129 L130 A131 P132 V133 V134 S135 M136 V137 R138 Q139 M140 Q141 E142 W143 T144 A145 P146 N147 E148 T149 R150 I151 Y152 F153 G154 V155 N156 T157 E158 P159 E160 L161 F162 Y163 I164 D165 E166 L167 K168 S169 L170 E171 R172 S173 M174 R175 N176 L177 T178 V179 K180 A181 C182 V183 W184 H185 P186 S187

G188
D189
W190
E191
G192
E193
Q194
G195
S196
P197
I198
L201
R202
E203
D204
L205
E206
S207
S208
D209
A210
N211
P212
D213
I214
Y215
L216
C217
G218
P219
P220
G221
M222
I223
D224
A225
A226
C227
E228
L229
V230
R233
G234
I235
P236
Q239
V240
F241
F242
E243
K244
F245
L246
P247
S248
G249
A250
A251

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing*.

Of the 50 calculated structures, 10 were deposited, based on the following criterion: *target function*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
DYANA	structure solution	1.5
X-PLOR	refinement	3.84

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	BMRB entry 6295
Number of chemical shift lists	1
Total number of shifts	2550
Number of shifts mapped to atoms	2544
Number of unparsed shifts	0
Number of shifts with mapping errors	6
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	75%

No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality (i)

6.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: FDA

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 (average) 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	1.53±3.72	0±1/1870 (0.0±0.0%)	0.56±0.44	1±2/2537 (0.0±0.1%)
All	All	4.02	3/18700 (0.0%)	0.71	6/25370 (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	Planarity
1	A	0.0±0.0	0.3±0.9
All	All	0	3

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	55	ILE	CA-CB	368.85	10.03	1.54	6	1
1	A	55	ILE	N-CA	311.24	7.68	1.46	6	1
1	A	55	ILE	CA-C	260.64	8.30	1.52	6	1

5 of 6 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	55	ILE	CB-CA-C	-49.91	11.77	111.60	6	1
1	A	55	ILE	N-CA-CB	-45.85	5.34	110.80	6	1
1	A	54	THR	C-N-CA	-44.22	11.14	121.70	6	1
1	A	55	ILE	N-CA-C	-34.79	17.06	111.00	6	1
1	A	55	ILE	CA-C-O	-24.43	68.81	120.10	6	1

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	54	THR	Peptide	1
1	A	55	ILE	Mainchain,Peptide	1

6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1825	1785	1784	472±32
2	A	53	33	33	19±5
All	All	18780	18180	18170	4731

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 128.

5 of 2339 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:155:VAL:HG21	1:A:161:LEU:HD13	1.12	1.17	10	8
1:A:53:LEU:HD12	1:A:102:LEU:HD11	1.12	1.22	1	7
1:A:31:LEU:HD22	1:A:102:LEU:HD11	1.10	1.20	7	2
1:A:155:VAL:HG21	1:A:161:LEU:HD22	1.06	1.28	3	4
1:A:79:LEU:HD12	1:A:131:ALA:HB2	1.06	1.28	2	5

6.3 Torsion angles (i)

6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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	233/250 (93%)	160±4 (68±2%)	41±4 (18±2%)	33±3 (14±1%)	1 5
All	All	2330/2500 (93%)	1595 (68%)	408 (18%)	327 (14%)	1 5

5 of 70 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	196	SER	10
1	A	155	VAL	10
1	A	210	ALA	10
1	A	195	GLY	10
1	A	58	THR	10

6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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	197/210 (94%)	108±5 (55±2%)	89±5 (45±2%)	0 2
All	All	1970/2100 (94%)	1078 (55%)	892 (45%)	0 2

5 of 162 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	163	TYR	10
1	A	54	THR	10
1	A	164	ILE	10
1	A	245	PHE	10
1	A	102	LEU	10

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains (i)

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

6.5 Carbohydrates (i)

There are no carbohydrates in this entry.

6.6 Ligand geometry (i)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds 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 is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	FDA	A	252	-	50,58,58	2.05±0.02	1±0 (1±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of 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 angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	FDA	A	252	-	50,89,89	1.80±0.02	1±0 (1±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
2	FDA	A	252	-	-	0±0,30,50,50	0±0,6,6,6

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	A	252	FDA	C9A-N10	5.32	1.46	1.38	5	8

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	A	252	FDA	C4-C4X-C10	5.59	117.95	114.57	4	9

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 75% for the well-defined parts and 75% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 6295

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	2550
Number of shifts mapped to atoms	2544
Number of unparsed shifts	0
Number of shifts with mapping errors	6
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	7

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atoms in chemical component dictionary. First 5 (of 6) occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	252	FDA	1HM7	8.506	-1.0	1
A	252	FDA	2HM8	6.899	-1.0	1
A	252	FDA	1HM8	6.711	-1.0	1
A	252	FDA	AH61	7.163	-1.0	1
A	252	FDA	AH62	7.985	-1.0	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	250	-0.41 ± 0.11	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	217	0.54 ± 0.06	Should be applied

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C'$	221	-0.64 ± 0.20	Should be applied
^{15}N	231	-0.53 ± 0.32	None needed (imprecise)

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 75%, i.e. 2137 atoms were assigned a chemical shift out of a possible 2853. 42 out of 42 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	1081/1134 (95%)	426/450 (95%)	439/468 (94%)	216/216 (100%)
Sidechain	987/1498 (66%)	546/884 (62%)	430/542 (79%)	11/72 (15%)
Aromatic	69/221 (31%)	66/117 (56%)	0/98 (0%)	3/6 (50%)
Overall	2137/2853 (75%)	1038/1451 (72%)	869/1108 (78%)	230/294 (78%)

7.1.4 Statistically unusual chemical shifts (i)

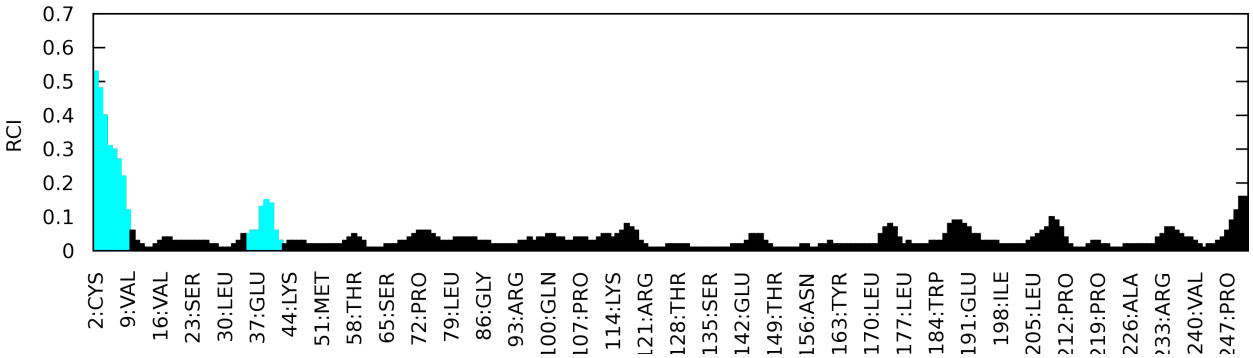
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	139	GLN	HG3	0.60	3.75 – 0.85	-5.9
1	A	26	THR	HG23	-0.16	2.29 – -0.01	-5.7
1	A	26	THR	HG22	-0.16	2.29 – -0.01	-5.7
1	A	26	THR	HG21	-0.16	2.29 – -0.01	-5.7
1	A	139	GLN	HB2	0.65	3.30 – 0.80	-5.6
1	A	88	PHE	H	12.03	11.96 – 4.76	5.1
1	A	108	LEU	HB2	-0.09	3.32 – -0.08	-5.0

7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:


RCI
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2:CYS
9:VAL
16:VAL
23:SER
30:LEU
37:GLU
44:LYS
51:MET
58:THR
65:SER
72:PRO
79:LEU
86:GLY
93:ARG
100:GLN
107:PRO
114:LYS
121:ARG
128:THR
135:SER
142:GLU
149:THR
156:ASN
163:TYR
170:LEU
177:LEU
184:TRP
191:GLU
198:ILE
205:LEU
212:PRO
219:PRO
226:ALA
233:ARG
240:VAL
247:PRO