

September 2004

^1H , ^{13}C , and ^{15}N assignments for the *Archaeoglobus fulgidis* protein AF2095

Robert Powers

University of Nebraska - Lincoln, rpowers3@unl.edu

Thomas B. Acton

Rutgers University, acton@cabm.rutgers.edu

Yi-wen Chiang

Rutgers University, chiang@cabm.rutgers.edu

P. K. Rajan

Rutgers University

John R. Cort

Pacific Northwest National Laboratory, Richland, Washington

See next page for additional authors

Follow this and additional works at: <http://digitalcommons.unl.edu/chemistrypowers>

 Part of the [Chemistry Commons](#)

Powers, Robert; Acton, Thomas B.; Chiang, Yi-wen; Rajan, P. K.; Cort, John R.; Kennedy, Michael A.; Liu, Jinfeng; Ma, LiChung; Rost, Burkhard; and Montelione, Gaetano T., " ^1H , ^{13}C , and ^{15}N assignments for the *Archaeoglobus fulgidis* protein AF2095" (2004). *Robert Powers Publications*. 3.
<http://digitalcommons.unl.edu/chemistrypowers/3>

This Article is brought to you for free and open access by the Published Research - Department of Chemistry at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Robert Powers Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

Robert Powers, Thomas B. Acton, Yi-wen Chiang, P. K. Rajan, John R. Cort, Michael A. Kennedy, Jinfeng Liu, LiChung Ma, Burkhard Rost, and Gaetano T. Montelione

LETTER TO THE EDITOR: ^1H , ^{13}C , and ^{15}N assignments for the
Archaeoglobus fulgidis protein AF2095

Robert Powers^{a,*}, Thomas B. Acton^b, Yiwen Chiang^b, P. K. Rajan^b,
John R. Cort^c, Michael A. Kennedy^c, Jinfeng Liu^d, LiChung Ma^b,
Burkhard Rost^d & Gaetano T. Montelione^{b,e}

^aDepartment of Chemistry, University of Nebraska–Lincoln, Lincoln, NE 68588

^bCenter for Advanced Biotechnology and Medicine, Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, NJ 08854

^cBiological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington, DC 99352.

^dDepartment of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032

^eDepartment of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, NJ 08854.

*To whom correspondence should be addressed. E-mail: rpowers3@unl.edu

Submitted March 2004; accepted May 4, 2004.

Key words: *Archaeoglobus fulgidis*, NMR, protein AF2095, resonance assignments, secondary structure

Biological context

Structural genomics is providing a means to determine the molecular and cellular function for the vast amount of proteins in the Human proteome that lack any explicit experimental information by characterizing the complete range of protein folds (Montelione, 2001). The Northeast Structural Genomics Consortium (NESG; <http://www.nesg.org/>) is a pilot project funded by the National Institutes of Health Protein Structure Initiative, focusing on proteins from eukaryotic model organisms including humans. The thermophilic archaea *Archaeoglobus fulgidis* AF2095 protein is an example of a protein of unknown biological function targeted for structural analysis by NESG. AF2095 belongs to the Pfam family PF01981 – UPF0099, protein domain family of unknown function that has been found in yeast, archaeobacteria and eubacteria. AF2095 has been assigned to NESG Cluster ID:17431, a set of fourteen proteins with high (>~30%) sequence identity with human, *Drosophila*, *Caenorhabditis elegans*, *Arabidopsis*, yeast, archaeal and eubacterial origin (Liu, 2004). A total of fifty-six proteins are identified when the analysis is expanded to include all available genomes, where determining the NMR solution structure of AF2095 can be leveraged to infer 3D structural information for these proteins. Here we report the near complete ^1H , ^{15}N , and ^{13}C NMR assignments and secondary structure of AF2095. These data provide a basis for determining the solution structure of AF2095, for further investigation of the function of this protein and for pro-

viding representative structural and functional information for the protein domain family that includes AF2095.

Methods and experiments

Uniformly ^{13}C , ^{15}N -enriched AF2095 (123 amino acids) was cloned, expressed and purified following standard protocols used in the NESG consortium. Briefly, the full length gene (YK95_ARCFU) from *Archaeoglobus fulgidis* was cloned into a pET21d (Novagen) derivative, yielding the plasmid pGR4-21. The resulting AF2095 open reading frame contains eight nonnative residues at the C-terminus (LEHHHHHH) of the protein. *Escherichia coli* BL21 (DE3) pMGK cells, a rare codon enhanced strain, were transformed with pGR4-21, and cultured in MJ9 minimal medium (Jansson *et al.*, 1996) containing $(^{15}\text{NH}_4)_2\text{SO}_4$ and $U\text{-}^{13}\text{C}$ -glucose as sole nitrogen and carbon sources. Initial growth was carried out at 37° C until the OD₆₀₀ of the culture reached ~ 0.8 units. The incubation temperature was then decreased to 17°C and protein expression was induced by the addition of IPTG (isopropyl-β-D-thiogalactopyranoside) at a final concentration of 1 mM. Following overnight incubation at 17°C, the cells were harvested by centrifugation and lysed by sonication. $U\text{-}^{13}\text{C}$, ^{15}N AF2095 was purified in a two step protocol consisting of Ni-NTA affinity column (Qiagen) and gel filtration column (Hi-Load 26/60 Superdex 75 pg, Amersham Biosciences) chromatography. The final yield of pure $U\text{-}^{13}\text{C}$, ^{15}N AF2095 (> 97% by SDS-PAGE; 13.5 kDa by MALDI-TOF mass spectrometry)

SUPPLEMENTARY MATERIAL for ^1H , ^{13}C and ^{15}N Assignments for *Archaeoglobus fulgidis* Protein AF2095. by

Robert Powers, Thomas B. Acton, Yiwen Chiang, Rajan Paranj, John R. Cort, Michael A. Kennedy, Jinfeng Liu,

LiChung Ma, Burkhard Rost and Gaetano T. Montelione

Table S1 ^{15}N , ^{13}C , ^{13}CO and ^1H resonance assignments for AF2095 at pH 6.5 and 40°C.^a

Residue	N	CO	C α	C β	Others
M1	-(-)	-	-(-)	-(-)	
T2	-(-)	179.7	61.7(4.15)	69.3(4.20)	C γ 21.5(1.33)
L3	126.5(9.39)	173.9	54.2(4.93)	44.7(1.04,1.94)	C γ 27.3(1.83);C δ 27.1(0.88);C δ 23.5(0.69)
K4	117.1(8.95)	178.7	55.5(4.91)	37.2(1.72,1.56)	C γ 24.1(1.17,1.06);C δ 30.0(1.57); C ϵ 41.1(2.63)
Q5	123.1(8.42)	176.4	54.4(4.68)	35.1(1.82,1.73)	C γ 35.5(1.58);N ϵ -(7.36,5.29)
V6	125.8(8.80)	176.9	60.6(4.92)	34.7(1.70)	C γ 21.5(0.73);C γ 22.2(0.82)
I7	128.0(8.94)	176.9	60.9(4.79)	40.7(1.67)	C γ 30.3(1.32,1.13);C γ_m 20.9(0.73); C δ 15.0(0.83)
V8	127.5(9.16)	175.4	60.1(4.68)	32.3(1.92)	C γ 20.9(0.68);C γ 23.8(0.76)
V9	119.7(8.79)	175.3	59.0(4.94)	34.4(2.10)	C γ 19.2(0.78);C γ 21.2(0.83)
R10	119.5(7.68)	174.7	55.5(4.36)	32.8(1.62)	C γ 26.4(1.74);C δ 44.4(3.26)
D11	122.1(9.04)	173.5	54.3(4.89)	43.4(2.55,2.48)	
D12	115.4(8.88)	174.2	54.7(4.40)	39.1(2.81)	
L13	118.0(6.80)	-	54.2(4.31)	43.0(1.31,1.67)	C γ 27.4(1.50);C δ 22.8(0.76)
K14	119.9(8.20)	-	55.6(4.07)	30.0(1.81)	C γ 24.8(-);C δ 29.4(-);C ϵ 42.4(-)
L15	120.0(7.69)	-	53.9(4.48)	43.8(1.29)	
S16	-(-)	175.8	57.8(4.44)	64.3(3.22,3.13)	
R17	120.5(8.83)	171.9	60.5(3.91)	30.7(1.84)	C γ 27.4(1.61);C δ 43.7(3.23)
G18	106.5(8.67)	175.1	47.0(3.63,4.03)		
K19	120.8(7.37)	170.7	58.6(4.10)	32.5(2.00,1.81)	C γ 25.8(1.54,1.48);C δ 28.4(1.71); C ϵ 42.4(3.00)
L20	121.5(8.81)	172.0	58.6(3.86)	41.4(1.84,1.54)	C γ 27.1(1.57);C δ 25.1(0.84);C δ 24.1(0.84)
A21	119.3(7.76)	171.9	55.5(3.76)	17.5(1.45)	
V22	115.3(7.41)	174.1	67.5(3.08)	31.2(2.16)	C γ 22.4(0.57);C γ 24.2(0.98)
Q23	116.0(7.62)	172.0	58.1(3.76)	27.6(2.08)	C γ 33.6(2.34);N ϵ -(6.43,5.51)
V24	117.6(7.80)	173.7	65.9(3.38)	31.0(2.16)	C γ 20.9(0.69);C γ 22.8(0.85)
A25	122.6(7.97)	171.4	55.4(3.96)	18.6(1.27)	
H26	116.3(8.68)	171.8	56.9(4.35)	31.5(3.40,2.77)	C δ_2 116.7(6.69)
A27	119.5(7.85)	172.5	54.7(3.46)	19.9(1.35)	
A28	119.4(8.31)	169.7	54.6(4.10)	19.4(1.53)	
I29	117.6(7.53)	173.0	63.8(3.55)	36.1(2.47)	C γ 28.4(1.70,1.39);C γ_m 16.6(0.52);

I30	119.4(7.94)	172.2	65.7(3.54)	37.8(1.76)	C δ 11.2(0.83) C γ 29.3(1.50,1.09);C γ_m 16.6(0.69); C δ 13.0(0.78)
G31	104.5(8.47)	174.7	47.3(3.48)		
Y32	122.6(7.98)	172.9	63.4(3.50)	38.7(3.37,3.03)	C δ 132.6(7.05);C ϵ 118.1(6.86)
L33	118.7(8.63)	171.4	57.5(3.76)	42.2(1.88,1.46)	C γ 26.9(1.89);C δ 26.4(0.80);C δ 22.5(0.90)
K34	115.1(7.39)	174.0	56.9(4.14)	33.8(1.88,1.49)	C γ 25.1(1.40);C δ 28.7(1.70);C ϵ 42.1(3.02)
S35	114.8(7.08)	178.1	59.9(4.35)	64.8(3.73,3.54)	
D36	122.9(9.01)	173.3	54.9(4.35)	43.4(2.80,2.58)	
S37	119.9(8.59)	173.8	61.8(3.85)	62.8(3.96)	
S38	118.6(8.32)	173.6	61.4(4.36)	62.4(4.01,3.93)	
L39	125.6(8.20)	170.9	58.6(4.09)	42.9(1.81,1.52)	C γ 27.9(1.65);C δ 25.4(0.82);C δ 25.7(0.98)
R40	116.2(8.56)	173.4	59.1(4.15)	28.5(2.00,1.37)	C γ 25.4(1.69);C δ 43.4(3.00)
R41	121.8(7.25)	172.4	59.1(4.34)	30.2(2.08)	C γ 26.7(1.84,1.79);C δ 43.0(3.32)
K42	121.1(7.64)	172.0	59.4(4.15)	32.0(1.96)	C γ 24.8(1.51);C δ 29.0(1.72);C ϵ 42.1(3.00)
W43	118.6(8.44)	171.0	60.8(4.19)	27.6(3.85,2.93)	C δ_1 126.7(6.99);N ϵ_1 132.6(10.80) C ζ_3 120.6(7.90);C ζ_2 114.9(7.08); C η_2 122.5(7.14)
L44	121.2(8.16)	170.0	58.9(3.44)	41.9(2.14,1.57)	C γ 27.0(2.06);C δ 23.5(0.58);C δ 26.4(0.98)
D45	121.8(8.41)	172.5	57.3(4.35)	40.2(2.97,2.77)	
E46	116.8(7.62)	174.1	56.1(4.32)	30.4(2.44,2.11)	C γ 35.5(2.74,2.41)
G47	106.7(7.59)	176.6	45.0(3.46,4.21)		
Q48	119.7(8.30)	176.1	52.4(-)	28.8(-)	C γ 31.0(1.03)
K49	118.0(8.38)	171.7	58.8(4.09)	33.0(2.00)	C γ 25.4(0.20,-0.34);C δ 29.3(1.73); C ϵ 42.1(2.39)
K50	125.1(9.05)	175.6	54.7(6.01)	39.5(1.52,1.73)	C γ 23.8(1.24);C δ 30.6(1.57);C ϵ 41.4(2.63)
V51	119.0(8.25)	177.0	61.1(4.29)	36.1(1.92)	C γ 20.9(0.95)
V52	125.6(8.13)	175.6	60.8(5.33)	32.9(1.90)	C γ 22.5(0.85)
L53	128.6(8.99)	175.4	53.0(4.95)	45.7(1.59,1.37)	C γ 27.4(1.48);C δ 26.4(0.83);C δ 23.1(0.75)
K54	117.4(8.22)	174.9	54.5(5.51)	36.3(1.59)	C γ 23.8(1.24);C δ 29.7(1.57); C ϵ 41.7(2.85,2.96)
V55	111.7(8.12)	174.2	58.9(4.79)	34.9(2.43)	C γ 18.9(0.68);C γ 22.5(0.80)
K56	117.8(8.50)	174.9	57.8(4.35)	33.6(1.90)	C γ 24.4(1.50);C δ 29.0(1.70);C ϵ 41.7(3.00)
S57	107.6(7.14)	176.3	56.7(4.84)	67.0(4.25,3.92)	
L58	122.3(8.98)	172.6	57.7(3.96)	41.3(1.87,1.43)	C γ 26.7(1.54);C δ 25.4(0.99);C δ 23.1(0.84)
E59	118.3(8.77)	170.3	60.5(3.87)	28.7(2.08,1.92)	C γ 36.2(2.37,2.27)
E60	119.5(7.91)	171.4	59.4(4.06)	30.6(2.00)	C γ 36.8(2.34,2.27)
L61	121.4(7.62)	172.0	59.0(3.74)	42.4(2.09,1.54)	C γ 26.5(1.44);C δ 25.8(0.66);C δ 26.4(0.59)
L62	117.1(8.63)	170.4	57.1(3.91)	40.6(1.82,1.29)	C γ 26.4(1.76);C δ 25.1(0.85);C δ 21.2(0.71)
G63	109.2(8.14)	173.5	47.4(3.95)		
I64	123.6(7.77)	173.2	63.5(3.76)	36.6(2.09)	C γ 28.7(1.51,1.39);C γ_m 17.6(0.73); C δ 13.1(0.78)
K65	120.7(7.75)	172.7	60.5(3.69)	33.4(2.00,1.71)	C γ 24.8(1.22);C δ 30.0(1.57);C ϵ 41.7(2.90)
H66	115.3(8.33)	172.7	58.3(4.54)	28.3(3.36)	

K67	121.6(8.26)	171.1	59.4(4.03)	32.4(1.99,1.92)	C γ 25.1(1.46);C δ 29.0(1.74);C ϵ 42.1(2.95)
A68	120.8(8.43)	170.9	55.5(3.96)	18.3(1.37)	
E69	118.6(8.32)	169.3	59.4(4.15)	28.9(2.25,2.05)	C γ 36.5(2.55,2.44)
S70	117.6(8.09)	175.4	61.5(4.28)	62.8(4.01,3.92)	
L71	120.9(7.18)	173.4	54.6(4.43)	42.9(1.66)	C γ 27.1(1.70);C δ 23.1(0.83);C δ 25.8(0.79)
G72	107.4(7.86)	175.8	45.8(4.08,3.76)		
L73	119.9(7.43)	174.0	53.8(4.25)	42.8(1.43,1.13)	C γ 27.3(1.48);C δ 26.4(0.76);C δ 24.1(0.84)
V74	123.8(9.06)	174.8	64.1(3.75)	32.0(1.67)	C γ 21.5(1.00);C γ 22.8(0.72)
T75	116.6(7.78)	176.4	58.8(5.70)	73.4(3.96)	C γ 22.2(1.20)
G76	105.2(8.02)	179.3	44.6(4.41,3.30)		
L77	121.7(8.47)	173.5	54.0(4.98)	45.1(1.95,1.40)	C γ 27.4(1.71);C δ 24.8(0.80);C δ 26.4(0.80)
V78	121.2(8.15)	176.3	62.5(3.96)	33.0(2.05)	C γ 21.2(0.71);C γ 22.5(0.90)
Q79	124.8(7.83)	176.1	54.5(4.68)	32.5(1.82,1.73)	C γ 34.2(2.09,1.91);Ne -(6.73,7.31)
D80	123.5(9.19)	174.2	53.4(4.74)	43.3(2.88,2.74)	
A81	128.6(8.69)	171.8	53.7(4.25)	18.6(1.43)	
G82	105.4(8.79)	176.7	45.7(4.14,3.71)		
L83	122.7(7.63)	-	54.7(-)	41.7(-)	
T84	-(-)	-	-(-)	-(-)	
E85	-(-)	-	-(-)	-(-)	
V86	-(-)	-	-(-)	-(-)	
P87	-(-)	-	57.8(4.85)	32.3(2.29)	C γ 27.8(1.99);C δ 50.6(3.56,3.84)
P88	-(-)	172.9	64.0(4.28)	31.7(2.33,1.88)	C γ 27.4(2.10);C δ 50.9(3.62,3.96)
G89	112.1(8.89)	175.7	45.0(3.50,4.10)		
T90	118.0(7.40)	176.9	64.9(3.98)	69.6(3.82)	C γ 21.8(1.02)
I91	129.5(8.59)	174.2	60.6(4.57)	38.6(1.68)	C γ 27.4(1.80,1.55);C γ_m 18.9(0.83); C δ 13.4(0.91)
T92	116.4(8.56)	175.9	62.2(4.20)	69.5(3.99)	C γ 22.2(0.98)
A93	118.6(7.12)	175.3	51.4(5.28)	21.3(1.08)	
V94	117.3(8.81)	177.9	59.2(5.13)	36.6(1.76)	C γ 20.5(0.83);C γ 21.8(0.82)
V95	124.3(8.58)	178.4	58.0(5.18)	35.2(1.65)	C γ 16.0(0.80);C γ 22.2(0.80)
I96	125.5(9.02)	174.7	59.7(4.84)	41.2(1.54)	C γ 27.1(1.39);C γ_m 19.9(0.72);C δ 14.3(0.56)
G97	111.7(8.31)	-	43.4(4.15,3.35)		
P98	-(-)	175.7	60.2(4.26)	35.0(0.09,-0.04)	C γ 24.1(1.07,0.42);C δ 49.2(3.11,1.81)
D99	118.2(8.44)	175.9	52.7(4.40)	44.7(2.55,2.45)	
E100	123.0(11.15)	171.7	57.4(4.29)	30.2(2.17,1.84)	C γ 35.9(2.45,2.38)
E101	125.8(8.64)	-	60.9(3.71)	29.8(2.09,2.03)	C γ 36.2(2.21)
R102	116.2(8.85)	172.0	59.0(4.10)	29.9(1.86,1.93)	C γ 27.1(1.69);C δ 43.4(3.21)
K103	117.5(7.12)	172.6	59.1(4.06)	33.8(1.81)	C γ 25.4(1.47);C δ 29.7(1.65);C ϵ 42.1(3.00)
I104	117.7(7.64)	171.7	64.7(3.47)	37.9(1.62)	C γ 29.3(1.59,0.91);C γ_m 18.2(0.77); C δ 13.7(0.65)
D105	119.7(8.71)	172.0	56.6(4.34)	39.4(2.61)	
K106	118.6(7.16)	172.7	58.8(4.04)	32.5(1.96)	C γ 25.1(1.50);C δ 29.0(1.71);C ϵ 42.1(2.99)
V107	115.2(7.89)	173.4	64.0(3.90)	32.5(2.25)	C γ 22.5(0.89);C γ 19.6(0.79)
T108	106.4(7.77)	174.2	62.5(4.14)	68.4(4.39)	C γ 22.2(0.89)

G109	109.6(7.93)	176.4	46.8(3.88)		
N110	116.9(8.16)	175.1	52.8(4.84)	38.8(2.82,2.69)	N δ 112.2(7.41,6.74)
L111	122.1(7.72)	-	53.8(4.59)	41.6(1.84,1.49)	C γ -(1.71);C δ -(0.88)
P112	-(-)	174.6	62.4(4.87)	34.8(1.71)	C γ 24.4(1.96,1.78);C δ 50.2(3.58,3.47)
L113	122.7(8.64)	172.7	55.3(4.35)	42.8(1.59)	C γ 27.4(1.59);C δ 23.8(0.90);C δ 25.1(0.90)
L114	121.3(7.98)	173.8	56.0(4.17)	41.7(1.60)	C γ 27.1(1.40);C δ 24.4(0.78)
K115	122.3(7.90)	174.6	56.3(4.40)	33.6(1.76)	C γ 24.1(1.38);C δ 29.0(1.69);C ϵ 42.1(2.98)
L116	123.7(8.11)	173.6	55.3(4.29)	42.5(1.56)	C γ 26.7(1.56);C δ 23.5(0.99);C δ 24.8(0.81)
E117	122.2(8.24)	174.4	56.2(4.26)	30.7(1.86)	C γ 35.9(2.20)
H118	119.6(8.40)	176.9	55.3(4.67)	29.5(3.19,3.07)	
H119	125.4(8.24)	171.7	57.2(4.45)	29.7(3.22,3.13)	
H120	-(-)	-	-(-)	-(-)	
H121	-(-)	-	-(-)	-(-)	
H122	-(-)	-	-(-)	-(-)	
H123	-(-)	-	-(-)	-(-)	

Footnotes to Table S1

^aIn each column, ¹⁵N and ¹³C shifts are listed first, and the corresponding ¹H shifts are given in parentheses. ¹H, ¹³C and ¹⁵N chemical shifts are referenced according to the method of Wishart et al.

1995