Journal of Biomolecular NMR, **22:** 301–302, 2002.
KLUWER/ESCOM

Letter to the Editor: Base excision repair: NMR backbone assignments of *Escherichia coli* formamidopyrimidine-DNA glycosylase***

Garry W. Buchko[a], Susan S. Wallace[b] & Michael A. Kennedy[a,*]
[a]*Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352, U.S.A.;* [b]*Department of Microbiology and Molecular Genetics, The Markey Center for Molecular Genetics, University of Vermont, Burlington, VT 05405, U.S.A.*

Received 29 November 2001; Accepted 28 January 2002

Key words: chemical shift mapping, 7,8-dihydro-8-oxoguanine, DNA repair, Fpg (MutM) protein, oxidized DNA

Biological context

Oxidative DNA damage is one of the most important mechanisms responsible for mutagenesis, carcinogenesis, aging, and various diseases (Farr and Kogoma, 1991). To repair such damage virtually all aerobic organisms have a DNA repair mechanism, base excision repair (BER). Formamidopyrimidine-DNA glycosylase (Fpg), a 269-residue metalloprotein with a molecular weight of 30.2 kDa, is a key BER enzyme in prokaryotes. The principle biological role of Fpg is to remove the promutagenic DNA lesion 7,8-dihydro-8-oxoguanine (8-oxoG) (Boiteaux et al., 1987). Fpg binds double-stranded DNA and performs three catalytic activities: (i) DNA glycosylase, (ii) AP lyase, and (iii) deoxyribophosphodiesterase. In order to understand how Fpg performs its many biological activities on DNA in solution, the structure and function of *Escherichia coli* Fpg is being studied by NMR spectroscopy (Buchko et al., 2000).

Methods and experiments

Progress in making backbone assignments was accomplished using a combination of residue-specific labeling (Dahlquist et al., 1985) and perdeuteration (Venters et al., 1996) of the protein. Using 14 different residue-specific labeled samples the majority of the ^{15}N to ^{1}H^{N} cross peaks in the ^{15}N/^{1}H HSQC spectrum of Fpg were identified as arising from an A, C, G, E/Q, D/N, F, T, I, V, L, R, K, Y, or M. Using an $\sim$ 0.8 mM NMR sample of ^{2}H-, ^{13}C-, and ^{15}N-labeled Fpg, two-dimensional HSQC and three-dimensional HSQC-NOESY, HNCA, HNCOCA, HNCOCACB, HNCACB, HNCO, and HNCACO data** were collected at 25 °C using Varian 900-, 800-, 750- and 600-Niova spectrometers. In many of the experiments the data was collected with deuterium decoupling and TROSY.

Extent of assignment and data deposition

In total, cross peaks for 209 out of the 252 possible ^{1}H^{N} resonances (269 residues minus 16 prolines and the terminal amide) were observed (83%) and 180 of these were assigned (86%) (Figure 1A). Twenty-nine cross peaks labeled with a question mark could not be unambiguously assigned primarily because neither the 'i−1' ^{13}C$^{\beta}$ nor the 'i' ^{13}CO cross peak was observed in HNCOCACB or HNCACO experiments, respectively. Not shown in Figure 1A is the cross peak for G238, with a ^{1}H^{N} chemical shift of 3.03 ppm, representing the most shielded glycine ^{1}H^{N} reported in the BioMagResBank. Figure 1B shows the sequence of *E. coli* Fpg with the location of elements of secondary structure predicted on the basis of sequence alignment and crystal structure of *Thermus thermophilus* Fpg (Sugahara et al., 2000). The crystal structure of *T. thermophilus* Fpg shows the protein to consist of a C- and N-terminal domain with only ∼45% of the protein consisting of α-helices and β-sheets. The large number of turns and loops in the protein may account for the less than expected number of ^{15}N/^{1}H HSQC cross peaks. The residue-specific amide assignments

*To whom correspondence should be addressed. E-mail: makennedy@pnl.gov

**Pulse sequences for many of the experiments were obtained with permission from the public accessible ftp site of Lewis Kay at the University of Toronto, Canada.

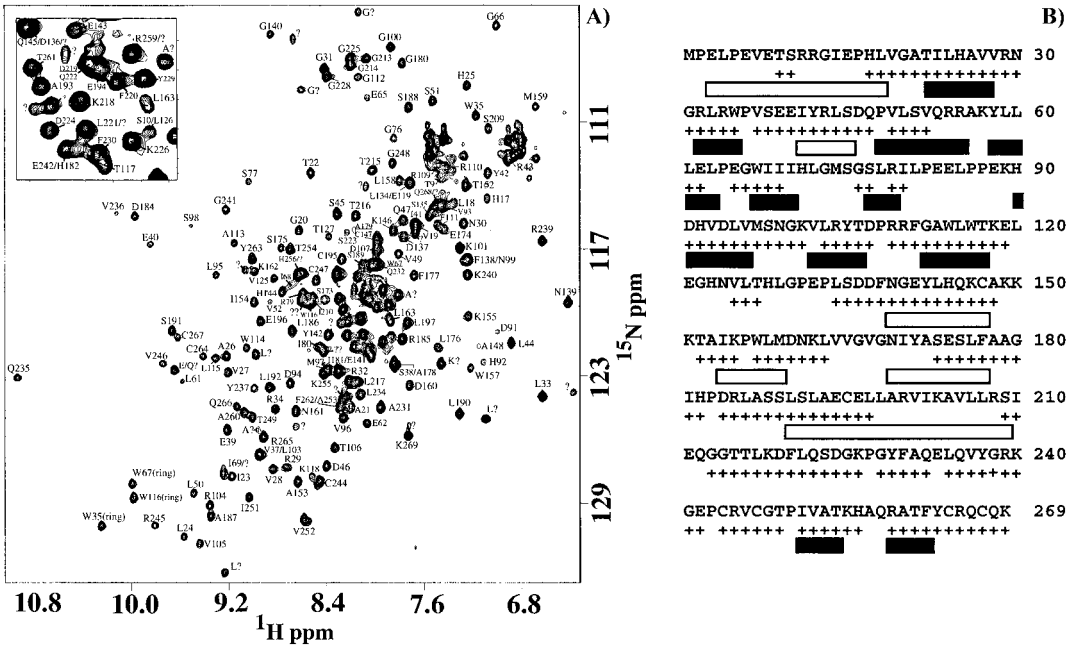

Figure 1. (A) Two-dimensional $^{15}N/^{1}H$ HSQC spectrum of Fpg ($\sim$ 0.8 mM, 300 mM potassium phosphate, 3 mM DTT, 50 μM NaN_3, pH 7.1, 25 °C) collected at a ^{1}H resonance frequency of 900 MHz. The residue-specific assignments of the cross peaks arising from the backbone amide groups are indicated. Backbone amide groups that could not be unambiguously assigned are labeled with a question mark (?). The remaining cross peaks correspond to the amide groups of asparagine and glutamine residues. The inset is an expansion of the congested region near the center of the spectrum. Not shown is residue G238 with a $^{1}H^{N}$ chemical shift of 3.03 ppm. (B) Amino acid sequence of *E. coli* Fpg. Amino acid residues highlighted with a plus (+) sign are residues for which $^{15}N/^{1}H$ HSQC cross peaks have been assigned. The predicted secondary structure for *E. coli* Fpg, based on sequence alignment and the crystal structure of *T. thermophilus* Fpg (Sugahara et al., 2000), is illustrated by open (α-helix) and closed (β-sheet) rectangles. The remainder of the structure is predicted to be composed of turns and loops.

for Fpg, indicated with plus (+) signs in Figure 1B, show that the assignments are distributed throughout the sequence with most of the assignments at, or nearby, residues predicted to adopt secondary structure. The only exceptions are a paucity of assigned $^{15}N/^{1}H$ HSQC cross peaks for residues in the first and last predicted α-helical regions which may indicate that these helical regions are shorter in *E. coli* than in *T. thermophilus* Fpg, or, these sites experience substantial motion in solution.

A table containing the ^{15}N, $^{1}H^{N}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$ and ^{13}CO chemical shifts are available from the authors as supplementary material and is deposited in the BioMagResBank in Madison WI, USA (accession number 5219).

Acknowledgements

The research was performed in the Environmental Molecular Sciences Laboratory (a national scientific user facility sponsored by the DOE Biological and Environmental Research) located at Pacific Northwest National Laboratory and operated for DOE by Battelle. This work was performed under the auspices of the U.S. Department of Energy (Contract DE-AC06-76RLO1830 for MAK and contracts DE-FG01-98ER45723 and DE-FG02-01ER64210 for SSW) and was supported by the DOE Office of Biological and Environmental Research Program under grant 24931 KP11-01-01. We thank Drs D.W. Hoyt and J.J. Ford for collecting data on the Varian-900 Niova spectrometer in Oxford, England.

References

Boiteux, S., O'Connor, T.R. and Laval, J. (1987) *EMBO J.*, **6**, 3177–3183.

Buchko, G.W., Hess, N.J., Bandaru, V., Wallace, S.S. and Kennedy, M.A. (2000) *Biochemistry*, **39**, 12441–12449.

Dahlquist, F.W., Griffey, R.H., McIntosh, L.P., Muchmore, D.C., Oas, T.G. and Redfield, A.G. (1985) In *Synthesis and Applications of Isotopically Labeled Compounds*, Muccino, R.R. (Ed.), Elsevier, Amsterdam, pp. 533–538.

Farr, S.B. and Kogoma, T. (1991) *Microbiol. Rev.*, **55**, 561–585.

Sugahara, M., Mikawa, T., Kumasaka, T., Yamamoto, M., Kato, R., Fukuyama, K., Inoue, Y. and Kuramitsu, S. (2000) *EMBO J.*, **19**, 3857–3869.

Venters, R.A., Farmer II, B.T., Fierke, C.A. and Spicer, L.D. (1996) *J. Mol. Biol.*, **264**, 1101–1116.