Characterisation of *Archaeglobus fulgidus* AlkA hypoxanthine DNA glycosylase activity

Colin Mansfield, Sinéad M. Kerins, Tommie V. McCarthy*

Department of Biochemistry, University College, Cork, Ireland

Received 16 February 2003; revised 7 March 2003; accepted 7 March 2003

First published online 18 March 2003

Edited by Ned Mantei

Abstract The AlkA protein from the archaebacterium *Archaeglobus fulgidus* was characterised with respect to release of hypoxanthine from DNA. The hypoxanthine glycosylase activity had optimal activity at 60°C at pH 5.0. The enzyme released hypoxanthine from substrates with a preference for dI:dG ≫ dI:dT > dI:dC > dI:dA. The presence of a mismatch on either side of the dIMP in the substrate reduced excision efficiency of the hypoxanthine residue at neutral pH, while a mismatch on both sides of the dIMP resulted in total loss of excision. Release of hypoxanthine from DNA required a minimum of two bases on the 5′ side and four bases on the 3′ side of the dIMP residue.

Key words: AlkA protein; Hypoxanthine; DNA glycosylase; DNA repair; *Archaeglobus fulgidus*

1. Introduction

The *Escherichia coli* AlkA (Ec-AlkA) protein is an inducible base excision repair DNA glycosylase that removes alkylation damage from DNA [1–3]. It releases a variety of modified bases from DNA including 3-MeAde, 3MeGua, 7-MeGua, 7-MeAde, O2-alkylpyrimidines, 1,N^6-ethenoadenine and 5-formyluracil [2,3]. In DNA, spontaneous deamination of dAMP generates 2′-deoxyinosine 5′-monophosphate (dIMP) [4]. Deamination of dAMP can also be induced by reactive oxygen species generated by ionising radiation and redox reactions [5]. dIMP residues can also occur in DNA by incorporation of dITP during replication [6]. The AlkA DNA glycosylase activity recognises dIMP and releases free hypoxanthine from DNA [7].

Hypoxanthine residues are mutagenic in cells if not excised as they behave like guanine with respect to base pairing and generate AT to GC transition mutations [8]. Removal of hypoxanthine is an activity that is found widespread and is associated in human cells with the ANPG protein [9], in rat cells with the APDG protein [10], and in yeast with the MAG protein [11]. When these glycosylases were compared with respect to hypoxanthine excision it was shown that ANPG and APDG proteins repair the hypoxanthine residues with an efficiency comparable to that of alkylated bases, while the *E. coli* AlkA and yeast MAG enzymes are quite inefficient [12].

The Ec-AlkA protein is a member of the Helix–hairpin–Helix (HhH) family of glycosylases that includes the endonuclease III, MutY, and 8-oxoguanine glycosylase proteins [13,14]. The crystal structure of AlkA revealed a globular protein consisting of three subdomains [13,15]. The conserved HhH motif is located at the interface of these subdomains, adjacent to the enzyme active site. AlkA utilises a base flipping mechanism to rotate the substrate base into the active site and the HhH motif anchors the DNA to the protein on the 3′ side of the flipped out nucleotide [16].

The purification and characterisation of the AlkA (methylpurine glycosylase) from the archaebacterium, *Archaeglobus fulgidus* was recently reported with respect to DNA glycosylase activity on methylated base substrates [17]. The purified AlkA protein differed from Ec-AlkA by excising alkylated bases only from DNA in the following order of efficiency: 3-methyladeneine ≫ 3-methylguanine ∼ 7-methyladeneine > 7-methylgluanine [17]. However, the glycosylase activity of the enzyme on hypoxanthine release from DNA was not investigated. Here we report characterisation of the hypoxanthine glycosylase activity of the *A. fulgidus* AlkA (Af-AlkA).

2. Materials and methods

2.1. Cloning, overexpression and purification of the Af-alkA gene

The *alkA* gene was amplified from *A. fulgidus* genomic DNA using primers designed from the published nucleotide sequence with added 5′ *Nde*I and *Sap*I restriction sites (5′-GACTTA**CATATGTGGAGAATTGAGCTGAA**-3′ and 5′-AAGCTC**GCTCTTC**CGCA**AACTAATTCTGTCT**-3′) (TIGR database entry AF2117) [18]. The *Nde*I and *Sap*I digested amplified product was cloned into *Nde*I and *Sap*I digested pTYB1 expression vector (New England Biolabs). The Ec-AlkA was cloned into the pTYB1 vector in a similar manner. The Ec-AlkA and Af-AlkA proteins were overexpressed and purified essentially as described for *Thermus thermophilus* UvrD [19] using a post IPTG induction period of 6 h.

2.2. Enzymatic assays

Oligonucleotides were 5′ end labelled using T4 polynucleotide kinase and [γ-^{32}P]ATP (3000 Ci/mmol Amersham) and annealed to an equimolar amount of complementary oligonucleotide. 300 ng of Ec-AlkA or Af-AlkA were incubated with 5 ng of each end labelled substrates in a total volume of 20 μl. Reactions were incubated for 60 min at either 37°C (Ec-AlkA) or 60°C (Af-AlkA). Following incubation, abasic sites were cleaved by adjusting the samples to a final concentration of 0.05 M NaOH and incubating at 95°C for 5 min. For determination of optimal pH, assays were performed in buffers (25 mM) of varying pH: sodium acetate/acetic acid, pH 5, morpholineethanesulfonic acid/NaOH, pH 5.5, 6.0 and 6.5, Tris–HCl, pH 7.0, 7.5, 8.0 and 8.5, and glycine/NaOH, pH 9.0, 9.5 and 10.0. Reactions were

*Corresponding author. Fax: (353)-21-4904259.
E-mail address: t.mccarthy@ucc.ie (T.V. McCarthy).

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)00257-6

stopped by transferring tubes to ice and adding an equal volume of loading solution (98% formamide, 10 mM EDTA, 0.5% xylene cyanol, and 0.5% bromophenol blue). The reaction products were analysed by denaturing gel electrophoresis (20% polyacrylamide/7 M urea) gels and quantified by phosphoimaging analysis using ImageQuant Software (Molecular Dynamics, Inc).

3. Results and discussion

3.1. Overexpression and purification of the A. fulgidus alkA gene

The open reading frame (ORF) of the Af-alkA gene was amplified from *A. fulgidus* genomic DNA using forward and reverse primers with incorporated *Nde*I and *Sap*I restriction sites respectively. PCR amplification of the *A. fulgidus* ORF of the *alkA* gene yielded the expected 911 bp fragment, which was subsequently digested and cloned into the pTYB1 expression vector as an *Nde*I/*Sap*I fragment. This placed the *alkA* gene under the control of the T7 promoter and directly upstream and in frame with the vector sequence encoding the intein–chitin binding domain. The 95 kDa AlkA fusion protein was induced in *E. coli* bearing the pTYB1-Af-AlkA expression plasmid. Following overexpression, the Af-AlkA protein was purified to >95% purity by chitin affinity chromatography from cell extracts. In this process, the C-terminal intein domain of the fusion protein was cleaved by incubation with dithiothreitol (DTT) resulting in specific elution of the native 40 kDa Af-AlkA (Fig. 1).

3.2. Hypoxanthine release DNA by Af-AlkA DNA glycosylase

The hypoxanthine DNA glycosylase activity of Af-AlkA was investigated and compared with the Ec-AlkA. Purified Af-AlkA or Ec-AlkA was incubated with duplex oligonucleotides bearing a central dIMP base paired with dA, dC, dG or dT. Following incubation, abasic sites were cleaved by treatment with NaOH. Glycosylase mediated cleavage of the substrates occurred with both enzymes. However, differences

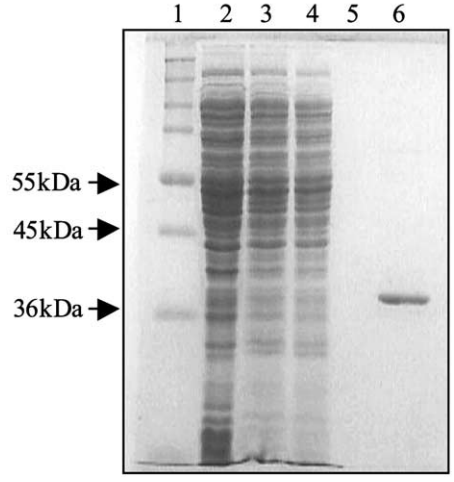


Fig. 1. SDS–PAGE (sodium dodecyl sulphate–polyacrylamide gel electrophoresis) (15%) analysis, of each stage of the purification of the Af-AlkA DNA glycosylase by chitin affinity chromatography. Lane 1: molecular weight markers; lanes 2 and 3: 13 μg of crude extract from *E. coli* pTYB1-Af-AlkA induced or not induced with 1 mM IPTG respectively; lane 4: 0.9 μg of proteins in extract from induced cells flowing through column; lane 5: 15 μl of wash buffer flowing through the column;lane 6: 0.3 μg of eluted fraction after DTT induced cleavage reaction.

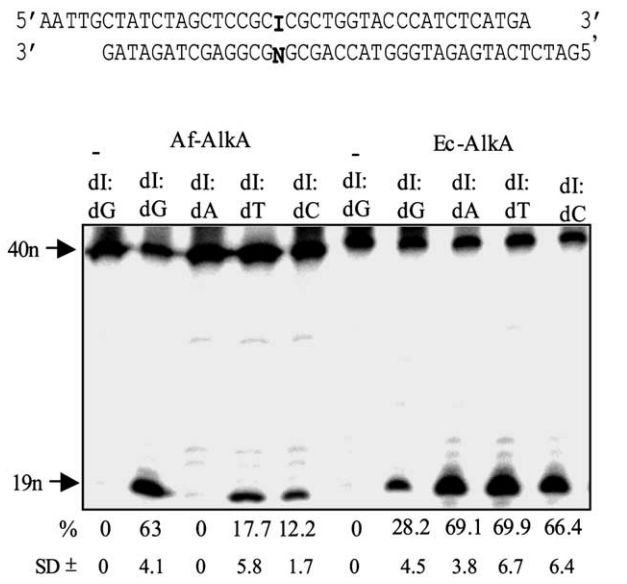


Fig. 2. Analysis of Af-AlkA and Ec-AlkA DNA glycosylase activity on duplex oligonucleotides with a central dIMP base paired with dG, dA, dT or dC. Each substrate was end labelled and incubated with 300 ng of Af-AlkA or Ec-AlkA glycosylase at 60°C or 37°C respectively in 25 mM HEPES–KOH buffer, pH 7.0. Lanes 1 and 6: dI:dG incubated at 60°C or 37°C respectively; lanes 2 and 7: dI:dG incubated with Af-AlkA and Ec-AlkA respectively; lanes 3 and 8: dI:dA incubated with Af-AlkA and Ec-AlkA respectively; lanes 4 and 9: dI:dT incubated with Af-AlkA and Ec-AlkA respectively; lanes 5 and 10: dI:dC incubated with Af-AlkA and Ec-AlkA respectively. % refers to % cleavage of substrate and is the mean of three experiments. − refers to the absence of protein. The % indicated is the mean of three experiments. Standard deviations are shown under each % cutting.

were observed with respect to substrate preference (Fig. 2). The Af-AlkA released the hypoxanthine base from the substrates with the following preference: dI:dG duplex (63.0 ± 4.1% release) ≫ dI:dT duplex (17.7 ± 5.8% release) > dI:dC duplex (12.2 ± 1.7% release) > dI:dA (0% release). The latter substrate was refractory to hypoxanthine release by the glycosylase. In contrast the substrate preference for the Ec-AlkA was dI:dT duplex (69.9 ± 6.7% release) ~ dI:dA duplex (69.1 ± 3.8% release) > dI:dC duplex (66.4 ± 6.4% release) ≫ dI:dG duplex (28.2 ± 4.5% release) (Fig. 2). It has previously been reported that the Ec-AlkA and the human equivalent (ANPG) has a substrate preference of dI:dT over dI:dG, dI:dC and dI:dA whereas the *Saccharomyces cerevisiae* equivalent (MAG) released hypoxanthine more efficiently from dI:dG over dI:dC, dI:dT and dI:dA [12]. Thus, the Af-AlkA is more similar to the MAG protein with respect to substrate preference. We have observed that altering the buffer conditions can change the efficiency of hypoxanthine release by the glycosylase, however, the relative activities of the AlkA enzymes on dI:dG, dI:dT, dI:dC and dI:dA did not vary (data not shown). The substrate preference for the Ec-AlkA observed here differs slightly from that reported with respect to activity on dI:dG, dI:dC and dI:dA [12]. This may reflect variation of activity with the sequence context of the substrate used. In support of this some differences in activity on dI:dG vs dI:dA have been reported with different substrates [12]. The Af-AlkA had little or no activity on hypoxanthine bases when present in single strand DNA (data not shown).

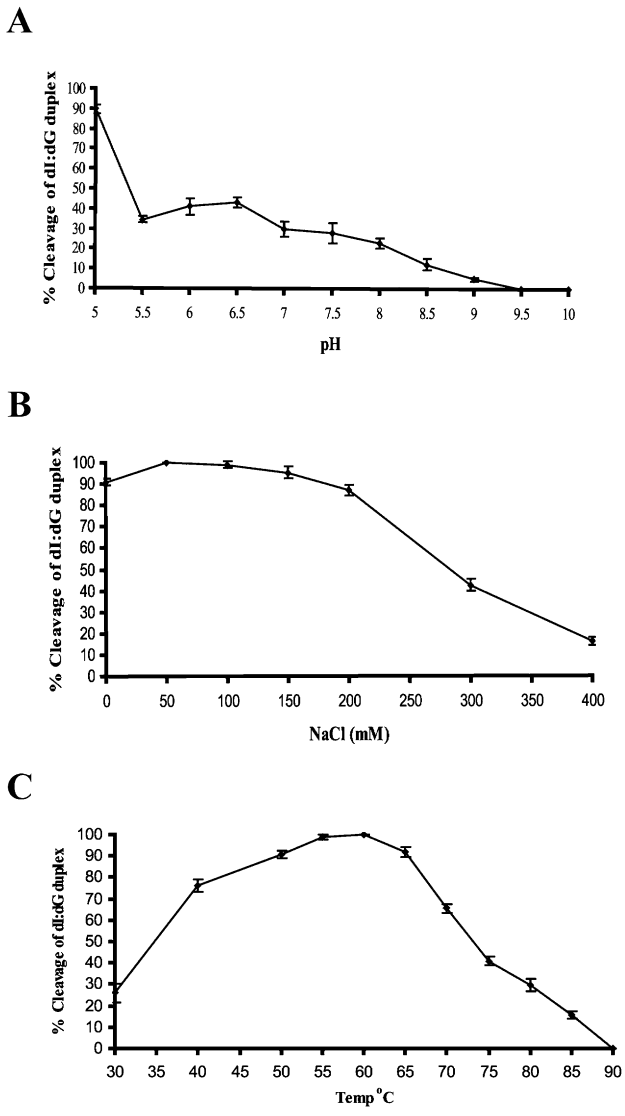


Fig. 3. The effect of pH (A), NaCl (B) and temperature (C) on the hypoxanthine DNA glycosylase activity of Af-AlkA. The 5′ end labelled dI:dG duplex substrate was incubated with 300 ng of the enzyme for 60 min under varying pH conditions (A) and in sodium acetate buffer pH 5 under varying NaCl concentration (B) and at varying temperatures and in the sodium acetate buffer, pH 5 (C). Experiments were carried out in triplicate and error bars show the standard deviation.

3.3. Characterisation of the Af-AlkA

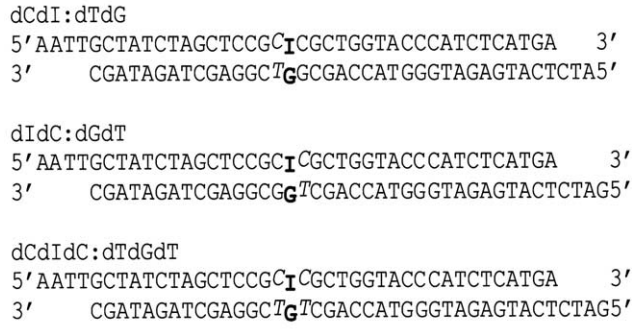
The pH optimum of the hypoxanthine DNA glycosylase activity of Af-AlkA was investigated over a pH range of 5.0–10.0. The hypoxanthine DNA glycosylase activity was active between pH 5.0 and 9.0. The highest activity of the enzyme was seen at pH 5.0 (Fig. 3A).

The effect of NaCl concentration on the hypoxanthine activity of Af-AlkA glycosylase was investigated at pH 5.0 over NaCl concentrations ranging from 0 to 400 mM. The Af-AlkA hypoxanthine glycosylase activity was sensitive to inhibition by NaCl at concentration above 150 mM (Fig. 3B). Inhibition of activity by NaCl at pH 7.5 was much more pronounced (data not shown).

The temperature optimum of Af-AlkA was investigated using the preferred dI:dG duplex substrate over a temperature range from 30°C to 90°C. Results showed that the Af-AlkA had an optimal temperature of 60°C (Fig. 3C) and worked efficiently between 50°C and 65°C with over 90% cutting of the substrate within these temperatures. This temperature profile of the hypoxanthine glycosylase activity of the enzyme is similar to the activity reported for release of methylated purines [17]. It was also observed that the activity of the Af-AlkA fell off rapidly after 65°C (Fig. 3C). The loss in activity is likely to be caused by generation of single strand regions in the substrate by local denaturation. The activity was stable at 70°C as preincubation of the Af-AlkA at 70°C for 60 min did not significantly reduce enzyme activity in subsequent assay at 60°C. Preincubation at 80°C for 60 min resulted in complete inactivation of the enzyme (data not shown).

3.4. The effect of mismatches on the hypoxanthine DNA glycosylase activity of Af-AlkA

The effect of mismatched flanking bases on AlkA DNA glycosylase activity has not been previously investigated. The influence of flanking mismatches on the AlkA hypoxanthine DNA glycosylase activity was investigated using a series of three duplex substrates identical in sequence to the dI:dG duplex substrate except for inclusion of flanking mismatches. The presence of a flanking 5′ C/T mismatch reduced the efficiency of the enzyme by 50% while a flanking C/T mismatch positioned on the 3′ side of the dIMP in the substrate reduced the efficiency by 65% (Fig. 4). When a 5′ and 3′ C/T mismatch flanked the dIMP, no hypoxanthine DNA glycosylase activity was detected (Fig. 4). Thus Af-AlkA hypoxanthine DNA gly-



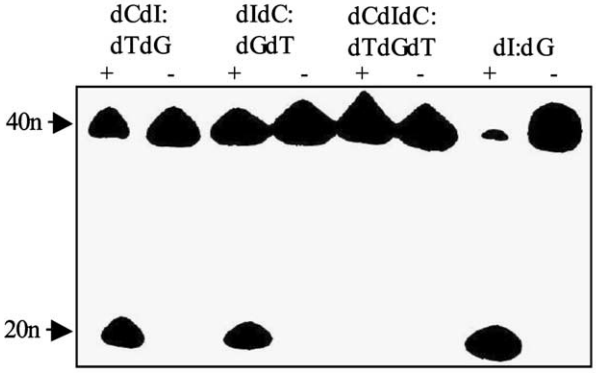


Fig. 4. Comparison of Af-AlkA DNA glycosylase activity on double stranded oligonucleotides, dCdI:dTdG, dIdC:dGdT, dCdIdC:dTdGdT, with the substrate dI:dG. Each substrate was incubated in 70 mM HEPES–KOH, pH 7.5, and 65°C for 60 min the presence and absence of Af-AlkA. + and − refer to the addition or absence of Af-AlkA respectively.

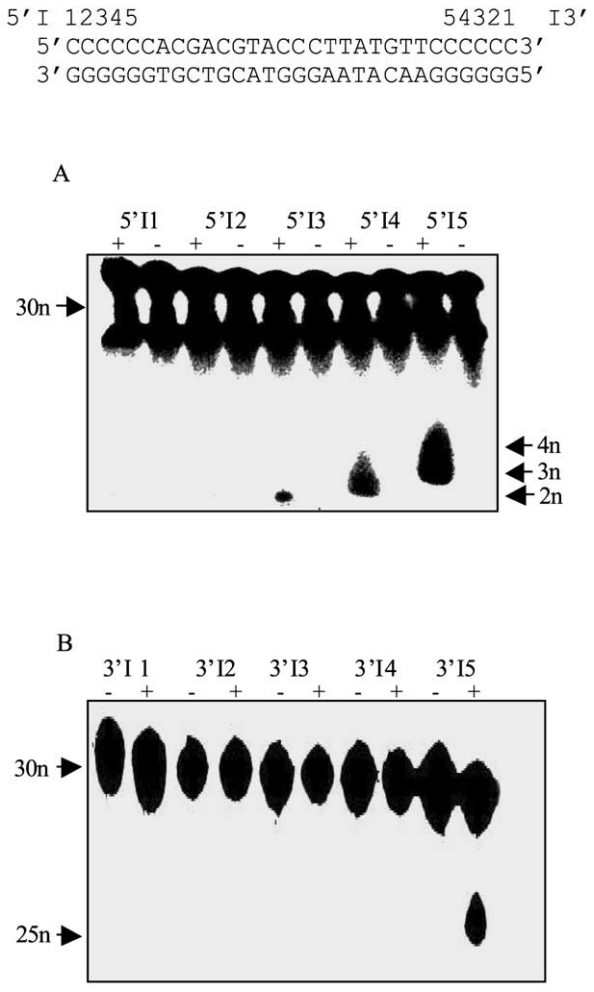


Fig. 5. Analysis of Af-AlkA DNA glycosylase release of hypoxanthine from 5′ and 3′ positions in a duplex substrate in 70 mM HEPES–KOH pH 7.5 and 65°C for 60 min. Duplex substrates were designed so that a single dI:dG base pair was present at 5′ position 1, 2, 3, 4 or 5 (5′I1, 5′I2, 5′I3, 5′I4, 5′I5) or 3′ position 1, 2, 3, 4 or 5 (3′I1, 3′I2, 3′I3, 3′I4, 3′I5). + and − refer to the addition or absence of Af-AlkA respectively.

cosylase activity is significantly reduced by mismatches flanking the dIMP target residue. The complete lack of activity of the enzyme on the target base flanked by mismatches is likely to reflect a strict double stranded substrate requirement for the enzyme at the target base region. This is in agreement with the known high resolution crystal structure of AlkA complexed to DNA which shows that there are about the same number of protein–DNA contacts on the DNA strand with the target base as on the complemenetary strand [20].

3.5. The activity of Af-AlkA DNA glycosylase on dI:dG base pairs in the 5′ and 3′ region of duplex substrates

The ability of Af-AlkA DNA glycosylase to release hypoxanthine residues from dI:dG base pairs in the 5′ and 3′ region of duplex substrates was investigated using double stranded oligonucleotides containing a single dI:dG base pair at the 5′ or 3′ end or at positions up to 5 nucleotides from the 5′ or 3′ ends. There was no release of hypoxanthine when dI was positioned at 5′ base position 1 or 2 while increasing amount of release occurred at 5′ positions 3, 4, and 5 respectively (Fig. 5A). By contrast, no release of hypoxanthine was observed when the dI:dG base pair was positioned at 3′ base position 1, 2, 3 or 4 while release was observed when the dI:dG base pair was at 3′ position 5 (Fig. 5B). AlkA is a member of HhH family of glycosylases. The crystal structure of AlkA shows that HhH motif anchors the DNA to the protein on the 3′ side of the target base [13,15]. These results show that in addition to anchoring the enzyme 3′ of the target base, a minimum of two bases 5′ of the target base are required for activity. Release of hypoxanthine was not observed when the base was positioned within four base pairs of the 3′ end. This indicates that a minimum of four bases is necessary for interaction with the HhH motif of the glycosylase.

To date, the AlkA protein of *A. fulgidus* is the only AlkA protein fully characterised from a thermophilic organism with respect to methylpurine DNA glycosylase activity [17] and hypoxanthine DNA glycosylase activity (this report). A new class of methylpurine DNA glycosylase, designated MpgII, has been described [21]. MpgII is a member of the endonuclease III family of DNA repair enzymes and has been purified from the thermophile *Thermotoga maritima*. The enzyme releases both 7-MeGua and 3-MeAde from DNA. However, it is not known if this enzyme has hypoxanthine DNA glycosylase activity.

Endonuclease V, a deoxyinosine specific endonuclease has also been characterised from *A. fulgidus*. This enzyme recognises hypoxanthine in DNA and cleaves DNA efficiently 3′ of the residue as the first step in a DNA repair pathway that has yet to be elucidated [22]. Endonuclease V is likely to be the main repair activity for hypoxanthine in DNA as the enzyme, unlike the Af-AlkA reported here, is highly active at inosine residues and cleaves DNA containing hypoxanthine to completion [22].

The substrate preference of Af-AlkA hypoxanthine DNA glycosylase activity is more similar to the *S. cerevisiae* enzyme than the *E. coli* enzyme. This may be explained by the fact that comparison of the protein sequence of the enzymes shows that the Af-AlkA is marginally more similar to *S. cerevisiae*. Similarly the conserved HhH motif on the protein is marginally more similar to *S. cerevisiae* [17].

Acknowledgements: This research was funded by BioResearch Ireland and the Higher Education Authority under the programme for research in third level institutions.

References

[1] Karran, P., Hjelmgren, T. and Lindahl, T. (1982) Nature 296, 770–773.
[2] Lindahl, T. (1982) Annu. Rev. Biochem. 51, 61–87.
[3] Sakumi, K. and Sekiguchi, M. (1990) Mutat. Res. 236, 161–172.
[4] Karran, P. and Lindahl, T. (1980) Biochemistry 19, 6005–6011.
[5] Wallace, S.S. (1983) Environ. Mutagen 5, 769–788.
[6] Mymes, B., Guddal, P.H. and Krokan, H. (1982) Nucleic Acids Res. 10, 3693–3701.
[7] Saparbaev, M. and Laval, J. (1994) Proc. Natl. Acad. Sci. USA 91, 5873–5877.
[8] Hill-Perkins, M., Jones, M.D. and Karran, P. (1986) Mutat. Res. 62, 153–163.
[9] O'Connor, T.R. and Laval, J. (1991) Biochem. Biophys. Res. Commun. 176, 1170–1177.
[10] O'Connor, T.R. and Laval, J. (1990) EMBO J. 9, 3337–3342.
[11] Berdal, K.G., Bjoras, M., Bjelland, S. and Seeberg, E. (1990) EMBO J. 9, 4563–4568.
[12] Saparbaev, M., Mani, J.C. and Laval, J. (2000) Nucleic Acids Res. 28, 1332–1339.

[13] Labahn, J., Scharer, O.D., Long, A., Ezaz-Nikpay, K., Verdine, G.L. and Ellenberger, T.E. (1996) Cell 86, 321–329.
[14] Guan, Y., Manue, R.C., Arvai, A.S., Parikh, S.S., Mol, C.D., Miller, J.H., Lloyd, S. and Tainer, J.A. (1998) Nat. Struct. Biol. 5, 1058–1064.
[15] Yamagata, Y., Kato, M., Odawara, K., Tokuno, Y., Nakashima, Y., Matsushima, N., Yasumura, K., Tomita, K., Ihara, K., Fujii, Y., Nakabeppu, Y., Sekiguchi, M. and Fujii, S. (1996) Cell 86, 311–319.
[16] Hollis, T., Ichikawa, Y. and Ellenberger, T. (2000) EMBO J. 19, 758–766.
[17] Birkeland, N.K., Anensen, H., Knaevelsrud, I., Kristoffersen, W., Bjoras, M., Robb, F.T., Klungland, A. and Bjelland, S. (2002) Biochemistry 41, 12697–12705.
[18] Institute for Genomic Research (online: http://www.tigr.org).
[19] Collins, R. and McCarthy, T.V. (2003) Extremophiles 7, 35–41.
[20] Hollis, T., Lau, A. and Ellenberger, T. (2000) Mutat. Res. 460, 201–210.
[21] Begley, T.J., Haas, B.J., Noel, J., Shekhtman, A., Williams, W.A. and Cunningham, R.P. (1999) Curr. Biol. 9, 653–656.
[22] Liu, J., He, B., Qing, H. and Kow, Y.W. (2000) Mutat. Res. 461, 169–177.