The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases

ABSTRACT

Here, using sequence profile searches, we show that several previously undetected protein families contain 2OG-Fe(II) oxygenase fold. This allows us to predict the catalytic activity for a wide range of biologically important, but biochemically uncharacterized proteins from eukaryotes and bacteria.

INTRODUCTION

Background 2-Oxoglutarate (2OG)- and Fe(II)-dependent dioxygenases are widespread in eukaryotes and bacteria and catalyze a variety of reactions typically involving the oxidation of an organic substrate using a dioxygen molecule. One well-studied reaction catalyzed by such enzymes is the hydroxylation of proline and lysine sidechains in collagen and other animal glycoproteins. In plants, enzymes of this family catalyze the formation of the plant hormone ethylene by oxidative desaturation of 1-aminocyclopropane-1-carboxylate, and catalyze the hydroxylation and desaturation steps in the synthesis of other plant hormones, pigments and metabolites such as gibberellins, anthocyanidins and flavones. In bacteria and fungi, several members of this family participate in the desaturative cyclization and oxidative ring expansion reactions in the biosynthesis of antibiotics such as penicillin and cephalosporin. The details of the catalytic mechanism of these enzymes have been revealed by determination of the crystal structures of isopenicillin N synthase (IPNS), deacetoxycephalosporin C synthase (DAOCS) and clavaminic acid synthase (CAS). These structures showed that the catalytic core of the proteins consists of a double-stranded β-helix (DSBH) fold containing a HX [DE] dyad (where X is any amino acid) and a conserved carboxy-terminal histidine which together chelate a single iron atom. The substrates are bound within a spacious cavity formed by the interior of the DSBH (see the Structural Classification of Proteins (SCOP)). We use sequence profile analysis to show that the DNA-repair protein AlkB, the extracellular matrix protein leprecan and the disease-resistance-related protein EGL-9 define new families of the 2OG-Fe(II) dioxygenase superfamily. AlkB is widely represented in bacteria and eukaryotes and has an important role in countering the toxic DNA modifications caused by alkylating agents in both Escherichia coli and Homo sapiens. Despite considerable effort, the precise biochemical mechanisms of its action in DNA repair remain unknown. Recent studies have shown that AlkB is required for specifically processing lesions resulting from the alkylation of single-stranded (ss) DNA. Our findings predict an unusual role for this enzyme in oxidative detoxification of DNA damage. The EGL-9 protein from Caenorhabditis elegans is necessary for normal muscle function, and its inactivation results in strong resistance to paralysis induced by the Pseudomonas aeruginosa toxin. We predict that EGL-9 is a novel hydroxylase that could elicit its action through the modification of sidechains of intracellular proteins. Similarly, we show that the animal extracellular matrix protein leprecan defines a hitherto unknown family of protein hydroxylases that might be involved in the generation of substrates for protein glycosylation.

CONCLUSION

Conclusions Before this study, structure determination, biochemical studies and sequence comparisons of 2OG-Fe(II) dioxygenases had elucidated their structural fold, active-site residues and reaction

mechanism. Here, using sequence profile searches, we show that many other protein families contain the same constellation of active-site residues and are predicted to adopt the same fold. This allows us to predict the catalytic activity of a wide range of functionally important, but biochemically uncharacterized, proteins from eukaryotes and bacteria. In particular, we propose a specific mechanism of action in DNA repair, and possibly in RNA modification, for the AlkB protein and its homologs.