Production of an active human AID enzyme in a bacterial system

LUCYNA BUDZKO [1], PAULINA JACKOWIAK [1], MAREK FIGLEROWICZ [1,2*]

[1] Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland

[2] Institute of Computing Science, Poznan University of Technology, Poznań, Poland

Polish Patent Application

P402440 (14.01.2013, Polish Patent Office)
Title of invention: *Modified gene of human activation-induced cytidine deaminase (AID), a method of modification of human AID gene, a composition showing AID activity, a method of preparation of such composition in a bacterial system, and use of the composition in the analyzes of DNA/RNA amination/deamination and/or methylation/demethylation*
Inventors: L. Budzko, P. Jackowiak, M. Figlerowicz
Assignee: Institute of Bioorganic Chemistry, PAS, Poznań
Application date: 14. 01. 2013

European Patent Application

PCT/PL2014/000003 (14.01.2014, PCT, WIPO)
Title of invention: *Modified gene of human activation-induced cytidine deaminase (AID), a method of modification of human AID gene, a composition showing AID activity, a method of preparation of such composition in a bacterial system, and use of the composition in the analyzes of DNA/RNA amination/deamination and/or methylation/demethylation*
Inventors: L. Budzko, P. Jackowiak, M. Figlerowicz
Assignee: Institute of Bioorganic Chemistry, PAS, Poznań
Application date: 14. 01. 2014

Abstract

The aim of our invention is to characterize a modified gene of human activation-induced cytidine deaminase (*hAID*), a method of modification of *hAID* gene, a composition showing hAID activity, and a method of preparation of such composition in a bacterial system. In particular, the invention concerns a modified *hAID* gene for the production of an active enzyme in a bacterial system as well as use of such composition for the analyzes of DNA/RNA amination/deamination, and/or methylation/demethylation.

Activation-induced cytidine deaminase (AID) is a key protein involved in the antibody diversification process. An unusual feature of the vertebrate immune system is the ability to produce large amounts of different antibodies, considerably more than the encoding genes. There are three mechanisms that increase the level of genetic variability of immunoglobulin genes. The primary mechanism is somatic recombination that occurs between V (*variable*), D (*diversity*), and J (*joining*) genes that encode the heavy and light chains of immunoglobulins. Such recombination allows to create a primary set of antibodies independent of the presence of an antigen. Following the B-lymphocytes activation by an antigen, secondary diversification of the immunoglobulin genes occurs by somatic hypermutation (SHM) and class switch recombination (CSR). The SHM is a process in which missense point mutations are introduced into the regions encoding antigen-binding domains of an antibody. CSR is a process that involves breaking of the DNA double-strand and recombination that induces chan-

ges in the effector (constant) region of the antibody heavy chain. As a result, the immunoglobulins that represent different classes (IgG, IgE, and IgA) are formed out of the primary M-type immunoglobulins (IgM) (Delker et al., 2009). The classes have different affinity for cell receptors and, therefore, they play different roles in activating the pathways for immune response. The molecular basis for SHM and CSR had been unknown until the end of the 20th century when the studies of T. Honjo's group led to the discovery of AID protein (Muramatsu et al., 1999) and showed that it is involved in the initiation of both of the above-mentioned processes (Muramatsu et al., 2000). In patients with mutated AID, severe immunological disorders were diagnosed: lack of somatic hypermutation, lack of class switch recombination (hyper IgM syndrome; HIGM2), and lymphoid hyperplasia (Revy et al., 2000). The role of AID in proper immune response is crucial; however, the up-to-date data show that the protein may play a more important role beyond the immune system. There are several evidences that show AID participates in genome demethylation. Now, it is known that genome methylation is one of the critical epigenetic factors that determine gene expression. Methylation of DNA occurs mainly at 5C-position of cytosine, which leads to the formation of 5-methylcytosine (5meC). In mammals, many enzymes have been identified that are responsible for both *de novo* methylation and maintaining methylation during DNA replication (Bestor, 2000). Very less is known about the molecular basis for the demethylation process. It may occur passively (where the methyl groups are not incorporated into the new DNA strands during replication) or actively (where methyl groups are removed from the DNA). Active, global genome demethylation occurs in mammalian development at two distinct stages: in the early embryo immediately after fertilization and in the primordial germ cells between 11.25 and 13.5 days after fertilization (Gehring et al., 2009). Demethylation restores a pluripotency state and leads to the reprogramming of the genome; thus, it is essential for the differentiation and development. The results of the latest studies suggest that AID may be involved in a gene-specific demethylation (Ramiro et al., 2015).

AID is a 24-kDa protein that belongs to the AID/APOBEC protein family. Members of this family represent a unique group of enzymes that function as DNA/RNA mutators. They can mutate DNA and/or RNA

as a result of their ability to deaminate cytidine to uridine. Currently, a crystal structure is only available for two members of the AID/APOBEC family. One of them is functionally uncharacterized APOBEC2 and the other is APOBEC3G (Prochnow et al., 2007; Conticello, 2008; Holden et al., 2008). However, for the latter only the structure of the carboxy-terminal domain has been established. The structure of AID remains undetermined and therefore, the exact catalytic mechanism of this enzyme is unknown. Similar to other proteins of the family, AID has a highly conserved amino acid motif that coordinates zinc atom inside the catalytic center. Therefore, it may be assumed that the mechanism of the cytidine to uridine deamination catalyzed by AID is similar to other zinc-dependent deaminases (Samaranayake et al., 2006).

Unfortunately, despite many recent efforts, the progress in AID research is far from satisfactory. One of the major reasons is the lack of an *in vitro* system which would enable the studies on molecular mechanisms of AID-mediated processes. Therefore, many questions concerning the natural activity and/or potential pathogenicity of AID remain open. For example, it is still not known how AID targets specific locations in the human genome. Because AID is a relatively small protein, one can hypothesize that to accomplish its functions, AID has to interact with other proteins. These proteins may regulate AID activity and selectively target it to specific genes or other regulatory elements present in the human genome. Another question is which cis-acting factors (encoded in AID structure) and trans-acting factors (encoded in DNA and other proteins' structure) determine the AID selectivity and site-specificity (Crouch et al., 2007). This issue is directly connected with the hypothesis that mutations identified in AID may be responsible for many serious immunological and developmental disorders and cancer. Thus, recognition of the molecular mechanisms behind AID activity seems to be of great importance for a better understanding of many physiological (e.g., immunoglobulin production; cell development, differentiation, and epigenetic reprogramming) and pathological (e.g., immunological defects, cancerogenesis) processes. At present, there is no doubt that AID is a very promising therapeutic target. Unfortunately, the lack of efficient *in vitro* system for testing AID activity hindered the search for AID inhibitors and modulators.

Several articles have been published on AID production: 1) in a bacterial system as a fusion protein with

glutathione S-transferase (GST) (Muramatsu et al., 1999; Papavasiliou et al., 2002) or with a streptavidin tag (Dickerson et al., 2003); 2) in insect cells (Bransteitter et al., 2004); and 3) in an *in vitro* translation system (Shivarov et al., 2008). The AID production in a bacterial system is the most efficient and economically effective, thus it is optimal in the production of large amounts of the enzyme for structural (including crystallographic) and biochemical studies (including the testing of the compounds that modulate the enzyme activity). Based on the literature, it may be assumed that the recombinant AID obtained so far in a bacterial system was contaminated with bacterial deaminase. As a result, in several publications, contradictory information appeared on AID functioning, among others concerning ability of AID to deaminate free nucleoside or inhibition of AID activity by tetrahydrouridine (both these features are characteristic for bacterial deaminases) (Beale et al., 2004; Samaranayake et al., 2006).

The aim of our study was to obtain a modified *hAID* gene that would enable the production of an active hAID enzyme in a bacterial system. A unique feature of this method is a modified *hAID* gene and a protein production process that involves a unique medium composition and optimized temperature conditions for the bacterial culture. In a process that did not involve the above-mentioned factors, the enzyme was produced inefficiently or did not adopt its active form. Notably, we excluded that the observed activity is derived from bacterial deaminase.

The proposed solution shall give new possibilities for thorough characterization of AID structure and biochemical activity. It will also enable to develop specific AID inhibitors that may be potentially applied to treat disorders caused by an increased level of AID expression, i.e., certain tumors and infections with hepatitis C virus or Epstein-Barr virus (Epeldegui et al., 2007; Machida et al., 2008; Marusawa, 2008). Owing to its high efficiency, the proposed method has already been applied in commercial production of hAID (Genesius, www. genesius.pl). Our commercialized invention has been nominated for the 2015 Marshal of Wielkopolska *i-Wielkopolska Innovative for Wielkopolska* Award. The invention which is the subject of the patent application shows a high level of innovation in two aspects: the process and the product. Process innovation derives from the fact that it is the first efficient method proposed to be used for producing a composition with hAID activity. Product

innovation involves the fact that the invention will allow the introduction of a new product that has not yet been available on the global market. The composition showing the activity of hAID (for methylated and unmethylated substrate) may be used as a potential tool for amination/deamination and methylation/demethylation research. Moreover, the composition can find buyers among scientists engaged in research into: 1) molecular mechanisms of genetic and infectious diseases, 2) immune processes, 3) mechanisms of carcinogenesis, 4) mechanisms of development, 5) differentiation and applications of stem cells, and 6) gene expression regulation. Besides, it can be used to develop and test compounds that modulate the AID activity (inhibitors and activators), including pharmacologically active substances. The high enzymatic activity, purity, and stability of the composition combined with low cost of its production will enable commercial success.

Patent claims included in the initial patent application:

1) A modified gene of human activation-induced cytidine deaminase (AID) characterized in that it comprises: 1) the entire cDNA sequence encoding human AID with STOP codons, except the 5′ and 3′ untranslated regions, including positions 80-679 in a sequence SEQ. ID No 1 (see Supplementary materials); 2) sequences encoding two tags: GST and hexahistidine; 3) sequences encoding two motifs recognized by the proteases: thrombin and enterokinase; and 4) sequences encoding elements of fusion protein in the following order (from the end 5′ to 3′): GST tag, a motif recognized by thrombin, hexahistidine tag, a motif recognized by enterokinase, AID; all being the sequence SEQ. ID No 3 (see Supplementary materials) or the sequences showing at least 85% amino acid similarity to SEQ. ID No 4 (see Supplementary materials).

2) A method for modification of the human activation-induced cytidine deaminase gene (AID) characterized in that it: 1) uses the entire cDNA sequence encoding human AID with STOP codons, except the 5′ and 3′ untranslated regions, including positions 80-679 in a sequence SEQ. ID No 1; 2) uses the sequences encoding two tags: GST and hexahistidine; 3) uses the sequences encoding two motifs recognized by the proteases: thrombin and enterokinase; and 4) assembles the sequences encoding elements of fusion protein in the following order (from the

end 5′ to 3′): GST tag, a motif recognized by thrombin, hexahistidine tag, a motif recognized by enterokinase, AID; all being the sequence SEQ. ID No 3 or the sequences showing at least 85% amino acid similarity to SEQ. ID No 4.

3) A composition showing AID activity for methylated and/or unmethylated cytidine characterized in that it contains a protein produced as a result of the expression of the modified gene in bacterial cells, according to claim 1.

4) A method of preparation (in a bacterial system) of the composition showing AID activity characterized in that it uses a modified gene of human AID, according to claim 1, wherein 1) uses the entire cDNA sequence encoding human AID with STOP codons, except the 5′ and 3′ untranslated regions, including positions 80-679 in a sequence SEQ. ID No 1; 2) uses the sequences encoding two tags: GST and hexahistidine; 3) uses the sequences encoding two motifs recognized by the proteases: thrombin and enterokinase; and 4) assemblies the sequences encoding elements of fusion protein in the following order (from the end 5′ to 3′): GST tag, a motif recognized by thrombin, hexahistidine tag, a motif recognized by enterokinase, AID; all being the sequence SEQ. ID No 3 or the sequences showing at least 85% amino acid similarity to SEQ. ID No 4.

5) A method according to claim 4 characterized in that it uses a modified gene of human activation-induced cytidine deaminase (AID) described in claim 1 and a bacterial culture is grown at the temperature from 15°C to 30°C, preferably 18°C.

6) A method according to claim 5 characterized in that a bacterial culture is grown on a medium that contains a source of zinc ions.

7) A method according to claim 6 characterized in that the source of zinc ions is zinc chloride at the concentration from 0.01 to 2 mM, preferably 0.06 mM.

8) Use of every protein, being a product of the expression of modified gene described in claim 1 and showing AID activity for methylated and/or unmethylated cytidine, in the analyzes of DNA/RNA amination/deamination and/or methylation/demethylation.

References

Beale R.C., Petersen-Mahrt S.K., Watt I.N., Harris R.S., Rada C., Neuberger M.S. (2004) *Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo*. J. Mol. Biol. 337: 585-596.

Bestor T.H. (2000) *The DNA methyltransferases of mammals*. Hum. Mol. Genet. 9: 2395-2402.

Bransteitter R., Pham P., Calabrese P., Goodman M.F. (2004) *Biochemical analysis of hypermutational targeting by wild type and mutant activation-induced cytidine deaminase*. J. Biol. Chem. 279: 51612-51621.

Conticello S.G. (2008) *The AID/APOBEC family of nucleic acid mutators*. Genome Biol. 9: 229.

Crouch E.E., Li Z., Takizawa M., Fichtner-Feigl S., Gourzi P., Montano C., Feigenbaum L., Wilson P., Janz S., Papavasiliou F. N., et al. (2007) *Regulation of AID expression in the immune response*. J. Exp. Med. 204: 1145-1156.

Delker R.K., Fugmann S.D., Papavasiliou F.N. (2009) *A coming-of-age story: activation-induced cytidine deaminase turns 10*. Nat. Immunol. 10: 1147-1153.

Dickerson S.K., Market E., Besmer E., Papavasiliou F.N. (2003) *AID mediates hypermutation by deaminating single stranded DNA*. J. Exp. Med. 197: 1291-1296.

Epeldegui M., Hung Y.P., Mcquay A., Ambinder R.F., Martinez-Maza O. (2007) *Infection of human B cells with Epstein-Barr virus results in the expression of somatic hypermutation-inducing molecules and in the accrual of oncogene mutations*. Mol. Immunol. 44: 934-942.

Gehring M., Reik W., Henikoff S. (2009) *DNA demethylation by DNA repair*. Trends Genet. 25: 82-90.

Holden L.G., Prochnow C., Chang Y.P., Bransteitter R., Chelico L., Sen U., Stevens R.C., Goodman M.F., Chen X.S. (2008) *Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications*. Nature 456: 121-124.

Machida K., Kondo Y., Huang J.Y., Chen Y.C., Cheng K.T., Keck Z., Foung S., Dubuisson J., Sung V.M., Lai M.M. (2008) *Hepatitis C virus (HCV)-induced immunoglobulin hypermutation reduces the affinity and neutralizing activities of antibodies against HCV envelope protein*. J. Virol. 82: 6711-6720.

Marusawa H. (2008) *Aberrant AID expression and human cancer development*. Int. J. Biochem. Cell Biol. 40: 1399-1402.

Muramatsu M., Kinoshita K., Fagarasan S., Yamada S., Shinkai Y., Honjo T. (2000) *Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme*. Cell 102: 553-563.

Muramatsu M., Sankaranand V.S., Anant S., Sugai M., Kinoshita K., Davidson N.O., Honjo T. (1999) *Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells.* J. Biol. Chem. 274: 18470-18476.

Papavasiliou F.N., Schatz D.G. (2002) *The activation-induced deaminase functions in a postcleavage step of the somatic hypermutation process.* J. Exp. Med. 195: 1193-1198.

Prochnow C., Bransteitter R., Klein M.G., Goodman M.F., Chen X.S. (2007) *The APOBEC-2 crystal structure and functional implications for the deaminase AID.* Nature 445: 447-451.

Ramiro A.R., Barreto V.M. (2015) *Activation-induced cytidine deaminase and active cytidine demethylation.* Trends Biochem. Sci. 40: 172-181.

Revy P., Muto T., Levy Y., Geissmann F., Plebani A., Sanal O., Catalan N., Forveille M., Dufourcq-Labelouse R., Gennery A., et al. (2000) *Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2).* Cell 102: 565-575.

Samaranayake M., Bujnicki J.M., Carpenter M., Bhagwat A.S. (2006) *Evaluation of molecular models for the affinity maturation of antibodies: roles of cytosine deamination by AID and DNA repair.* Chem. Rev. 106: 700-719.

Shivarov V., Shinkura R., Honjo T. (2008) *Dissociation of in vitro DNA deamination activity and physiological functions of AID mutants.* Proc. Natl. Acad. Sci. USA 105: 15866-15871.