1 Introduction

1.1 Colon carcinoma

1.1.1 Genetic alterations during colon carcinogenesis

Being one of the most frequent death reasons all over the world [1] colon cancer attracts attention of a large group of investigators. In 1990 Fearon and Fogelstein suggested a model of multistep development of colon carcinoma [2]. According to this model tumorigenesis proceeds through series of genetic alterations involving oncogenes [3] and tumour suppressor genes [4]. Adenomas developed from normal colonic epithelium proceed through different grades of dysplasia, and develop into carcinoma [2]. Invasive tumours continue to progress and the accumulation of further genetic alterations correlates with the ability of carcinoma to metastasise. This process requires years and possibly decades [2].

The typical tumour suppressor gene is p53, which is one of the common target for alteration in different human neoplasias [4]. In the case of sporadic colon cancer alterations of p53 chromosomal locus was observed in 70% of cases [5]. The protein product of this gene is responsible for the control of DNA damage, causing arrest in G1 phase of the cell cycle in order to allow DNA repair. In the case of the damage which can not be repaired p53 can induce apoptosis. Mutations of p53 usually occur at the later stages of progression of adenoma to carcinoma, i.e. the most of the early adenomas are free of p53 mutations [4]. Inactivating mutation of another gatekeeper – the adenomatous polyposis coli (APC) gene is assumed to be the earliest event in colorectal carcinogenesis [6]. Germline mutations of this gene are found in the condition of familial adenomatous polyposis, characterised by early development of colonic adenomas. This suggests that

APC has an important role in controlling the proliferation of colonic tissues. APC mutations occur in about 60 to 80% of sporadic colorectal carcinomas [7].

Mutations of Ki-ras oncogene are present in about 50% of colorectal cancers [2]. The frequency of mutations at later stages of colorectal cancer progression appeared identical to that found in large adenomas [2]. However several investigators have identified a number of Ki-ras gene mutations in 13% to 58% of aberrant crypt foci, which are putative precursors of colorectal cancer [8,9]. Thus Ki-ras gene activation appears to be an early event in colorectal carcinogenesis.

Studies of hereditary nonpolyposis colorectal cancer (HNPCC) revealed additional group of genes, classified as "caretakers" [10]. First gene identified was hMSH2 which belongs to a family of genes whose products are responsible for the correction of mistakes caused by the infidelity of DNA polymerase I during DNA replication [11]. Another part of the DNA repair complex - hMLH1 was identified in 1994 by Bronner et al [12]. Together these mutations in two genes account for approximately 60% of the HNPCC cases investigated. Additionally two other genes PMS1 and PMS2 were found to be mutated in some HNPCC patients [13]. This mutator genes themselves are not directly oncogenic, but they facilitate mutation of other genes that subsequently lead to neoplastic transformation.

1.1.2 Mucinous colon carcinoma

Phenotypic analysis of colon carcinomas allows to make a certain classification in accordance to the amount of the mucinous component in the tumour. Tumours with amount of mucinous component of at least 50% are defined as mucinous. Such tumours represent about 10 to 20 percent of all colon carcinomas. Mucinous carcinomas are less

frequent among sporadic colon carcinoma in contrast to HNPCC patients [14]. From the clinical point of view mucinous carcinomas of breast, ovary [15] and pancreas [16] have better prognosis than nonmucinous ones. In the case of colon carcinoma the data are controversial [17-19].

Differences in phenotype correlate with remarkable differences in genetical alterations. Mutations of p53 gene occur only in 30% of mucinous carcinoma [20,21], while the mutations of the protooncogene Ki-ras are more frequent in mucinous carcinomas (70%) than in non selected ones (50%) [14].

Another striking genetic alteration is the overexpression of the intestinal mucin gene MUC2 in 100% of mucinous carcinomas [22]. MUC2 gene is coding for a secretory mucin abundantly present in normal colon. Expression of MUC2 in nonmucinous carcinomas is significantly lower than in the normal colonic mucosa. The phenomenon of the overexpression of MUC2 gene becomes more interesting when aligned with the process of the colon carcinoma development. During the adenoma stage the significant increase of MUC2 gene product is observed in all cases. With further progression from adenoma to carcinoma the correlation of the expression of the gene with tumour phenotype is observed. In the case of nonmucinous tumours the expression is strongly decreased to the level below the one in normal tissue [22], and in the case of mucinous tumours the expression of MUC2 increases further (Fig. 1) [14,20].

Taken together these data suggest that mucinous and nonmucinous carcinomas develop through different genetic pathways and accumulate different sets of molecular lesions. Although the prognostic value of MUC2 overexpression is not clear, it is the only known molecular lesion 100% associated with mucinous phenotype. Therefore understanding of the mechanisms underlying the alterations of the expression of MUC2 gene can shed light on the general processes of colon carcinoma development.

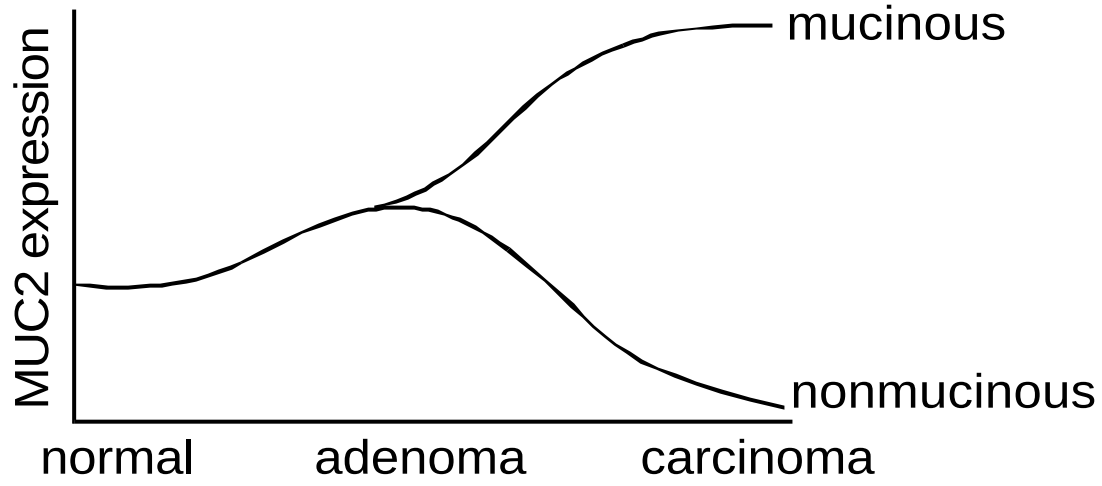

Figure1. *Alteration of MUC2 expression during colon carcinoma development.*

1.2 DNA methylation

Physiological DNA methylation - the only known covalent modification of DNA molecule - is accomplished by transfer of the methyl group from S-adenosyl methionin to 5 position of the purine ring of cytosine. DNA methylation is observed in most of the organisms at the different stages of evolution, in such a distinct species as *E.coli* and *H.sapience*. However some species, like *Drosophilae melanogaster* lack DNA methylation [23].

1.2.1 Function of DNA methylation in prokaryotes

The extensive research on methylation was conducted on bacteria. In this lower forms, both adenine and cytosine can be methylated, and this modification is involved in DNA replication and arrangement. A series of DNA methyltransferases (DNA-MTases) which can catalyse cytosine methylation in different sequence context were identified [24]. The main function of DNA methylation in bacteria is to provide a mechanism, which protects the cell from the effect of foreign DNA introduction. Restriction endonucleases discriminate between endogenous and foreign DNA by its methylation pattern. Introduced DNA which is not protected by methylation is then eliminated by cleavage [24].

Another function of DNA methylation in prokaryotes is the involvement in the control of replication fidelity. During DNA replication the newly synthesised strand does not get methylated immediately, but analysed for mismatches by the mismatch repair system. When a mutation is found the correction takes place on the nonmethylated strand [25].

1.2.2 Differences of DNA methylation between eukaryotes and prokaryotes

Eukaryotic DNA methylation affects only cytosine residues and specific for CpG sequence. However, the protective function of DNA methylation is similar in eukaryotes and prokaryotes. In humans and rodents inserted viral sequences can become methylated in association with silencing of the introduced genes [26]. The same mechanism is involved in silencing of transgenes in mice [27,28]. Thus function of DNA methylation machinery for recognition and/or eliminating of foreign DNA seem to be conserved in evolution.

The hypothesis on the involvement of DNA methylation in the repair process in eukaryotes was disproved by Araujo et al [29]. It was shown that methylation in

eukaryotic cells occurs immediately after replication and even Okazaki fragments are already methylated [29].

Due to the much higher complexity of eukaryotic genome in comparison to prokaryotic one it is logical to presume some additional roles of methylated cytosine as a "fifth base". Indeed, there is number of experimental evidences for the involvement of cytosine methylation in the functional reorganisation of eukaryotic genome. The regions of the genome with a high number of methylated cytosine are usually transcriptionally inactive. The absence of DNA methylation is a prerequisite for transcriptionally active regions. Since DNA methylation is reversible and does not directly depend on the sequence context it was described as an epigenetic mechanism of gene regulation [30,31].

1.2.3 Regulation of DNA methylation in eukaryotic cell

There are two basic types of normal methylation processes known in eukaryotic cells. First is *de novo* methylation which is involved in the rearrangement of methylation pattern during embryogenesis or differentiation processes in adult cells [32,33]. Recently a family of enzymes was described, containing two methyltransferases DNMT3a and DNMT3b which show the *de novo* methylation activity[34,35]. The homologous genes were identified in mouse [36]. Gene targeting experiments showed that both DNMT3a and DNMT3b are essential for *de novo* methylation and have no effect of maintenance methylation [37].

The second methylation activity in eukaryotic cell is the so-called maintenance methylation which is responsible for maintaining the methylation pattern once established. The first mouse maintenance methyltransferase DNMT1 was described by Bestor et al [38]. The enzymes with high homology were found in human [39] and chicken [40]. The functional analysis of the enzyme showed that it has maintenance

methylation activity and is vitally important for embryonic development in the mouse. The total homozygous knockout of mouse DNMT1 was lethal for the embryo [41,42]. During DNA replication DNMT1 is located in the replication complex where it recognises the normally methylated CpG sites in the parent strand and catalyses the addition of the methyl group in the corresponding CpG site in the daughter strand. Active localisation of the enzyme to sites of DNA replication in dividing cells may facilitate a maintenance role of DNMT1 [43]. One more methyltransferase – DNMT2 with unclear function was identified by Yoder and Bestor [44]. However already initial studies showed, that this enzyme is not essential for *de novo* methylation in eukaryotic cells [45].

To alter the established pattern of methylation there must be a mechanism responsible for the removal of existing methylation. There are two mechanisms known until now. First is a passive demethylation which occurs when DNMT1 fails to maintain the existing methylation pattern [46]. Second is active demethylation which is performed by recently described demethylase [47].

1.2.4 CpG islands

The distribution of CpG sites in the genome is as important as the role of DNMT1 activity. During the evolution, the sequence CpG has been progressively eliminated from the genome due to deamination of methylcytosines to thymines. For example in humans this dinucleotide is present only 5 to 10% of its predicted frequency. In 70 to 80% these CpG dinucleotides are methylated. These methylated regions are typical of the bulk chromatin that constitutes most nontranscribed DNA (for review [48])[49].

In contrast to the rest of the genome, smaller regions of DNA termed CpG islands, ranging in size from 0.5 to about 4 to 5 kb [49] have maintained the expected frequency of CpG content. A CpG island is defined as a sequence with a G+C content of greater

than 60% and ratio of CpG to GpC of at least 0.6 [50]. Most frequently these islands are located within 5' regulatory regions of genes. This may result from the fact that during evolution these regions were not methylated and, therefore, not depleted through the C to T transitions [48].

These two types of DNA methylation patterns determined by either low content of CpG sequence or CpG islands represent two types of regulatory regions. Genes which contain CpG island in their promoter are usually "housekeeping" genes, which have a broad tissue pattern of expression. Many relatively tissue specific genes are also regulated by CpG island methylation [48]. It is important to note that nonmethylated CpG island within the promoter region is not always associated with actively transcribed gene. However the lack of methylation of the CpG island within the promoter region of the gene is required for transcription of the gene. This modulatory role of methylation is reflected by the fact that chemically induced demethylation of CpG islands associated with inactivated genes leads to their partial reactivation [51].

CpG dinucleotides within the promoters without CpG island are usually methylated in a tissue-specific manner and can reflect the transcriptional status of the genes. In many cases these CpG sites are not methylated if the gene is actively expressed and methylated in cells with little or no transcription of the gene [49].

1.2.5 Regulation of transcription by methylation

In the case of CpG island-containing promoters the lack of methylation is usually associated with the chromatin pattern of actively transcribed genes, as characterised by an opened nucleosome configuration, reduces amount of histone H1 and presence of acetylated histones [52]. The ability of methylation to silence genes with CpG islands was studied on inactivated genes on X-chromosome [51]. Transfection studies showed, that

this silencing is mainly a result of chromatin condensation which makes DNA less accessible for transcription factors [53]. The role of single methyl groups, preventing binding of specific factors appears to be less important in this case.

In contrast to this, genes without CpG islands are dependent on the methylation of single sites within their promoter regions. This observation can be explained by the property of some transcription factors whose binding to DNA is methylation dependent, i.e. the protein binds to its binding site only in the case of nonmethylated DNA. Such methylation dependency was described for transcription factor AP-2 [54]. For another well characterised transcription factor - Sp1 the data are contradictory [55,56].

Another mechanism of transcription regulation by methylation of single sites involves methylation dependent binding proteins (MDBPs). MeCP1 described by Meehan et al. needs at least 7 methylated CpGs for efficient DNA binding and therefore is less important for genes without CpG islands in the promoters [57]. MeCP2 (MDBP-2) binds to a single methylated CpG and can inhibit the transcription of the gene [58]. In addition to this two proteins recently Hendrich and Bird described a family of MDBP which have high homology with MeCP2. All of them contain DNA binding domain as well as transcription inhibitory domain [59].

In summary methylation plays a very important role in the regulation of gene expression.

1.3 Role of DNA methylation in cancer

1.3.1 Methylation dependent mutations

CpG sites are hotspots for mutation in the human germline [60]. More recently it has become clear that they can be also hotspots for inactivating mutations in tumour suppresser genes [61,62]. About 25% of all mutations in p53 gene in all human cancers

studied occur at CpG sites, and almost 50% occur at methylation sites in colon cancer [63]. Since no endogenous chemicals have been found to increase directly the rate at which these mutations occur [64], they should be considered as part of an endogenous process.

The conventional explanation for the existence of the hotspots has been spontaneous hydrolytic deamination of 5-mCyt to T [65]. However, errors made during the methylation process may also contribute to mutagenesis. DNA methyltransferases can catalyse the deamination of C to U when S-adenosylmethionine is limiting [66]. Experiments of Yebra and Bhagwat [67] have shown that cytosine methyltransferases are also capable of the direct conversion of 5-mCyt to T, thus extending the repertoire of side reactions that could contribute to C->T transition.

The direct involvement of the DNA methyltransferase in mutagenesis at CpG sites would be expected to be facilitated by higher levels of enzyme expression, lower level of S-adenosylmethionine and decreased levels of specific repair enzymes. Although there is evidence for 4-3000 fold increase of methyltransferase activity in tumour cell lines [68], the existence of biochemical conditions favouring a C->U->T pathway in human colon tumours was not shown [69].

1.3.2 Overall decrease of DNA methylation in cancer cells

More than a decade ago it was shown that global genomic levels of DNA methylation are lower in cancer cells than in normal tissue [70-72]. In the number of experimental models of carcinogenesis, this decrease in numbers of methyl groups appears to begin early in tumour progression and before the tumour formation [73,74]. A possible direct role for DNA hypomethylation in the neoplastic process has been proposed from experimental

data showing that in rodents depletion of methyl donor from the diet results in liver carcinogenesis and in DNA hypomethylation [75].

Despite the clear association of DNA hypomethylation with both spontaneous and experimentally derived tumours, the exact role of this change is poorly understood. In 1983 Feinberg and Vogelstein reported a decrease of methylation in the promoter regions of c-Ha-ras and c-Ki-ras in lung and colon carcinomas [76]. Therefore activation of oncogenes was proposed as a possible role of decrease in DNA methylation in carcinogenesis. However no significant data was collected to test this hypothesis.

Schmidt et al. observed the abnormalities in chromosomal division during cell replication after decrease of overall methylation induced by 5-aza-2-deoxycytidine treatment [77]. This suggests that demethylation may influence the structural integrity of chromosomes leading to cell transformation [78]. However more studies are required to establish the consequence of DNA demethylation in neoplastic cells.

1.3.3 Regional hypermethylation in cancer

The same tumour cells which were described to have the overall genomic hypomethylation frequently have regions of dense hypermethylation. The fact that most of nonmethylated cytosines are located within CpG islands suggests that the normally nonmethylated CpG islands within 5' regulatory regions are the primary targets for aberrant hypermethylation in tumour cells [49].

Baylin et al described that a CpG island in the promoter region of the calcitonin gene at chromosome 11p, which was unmethylated in all normal tissues tested, was densely methylated in human solid tumours [79], leukemias [80] and cells transformed with various viruses [80,81]. Additionally some other CpG-rich regions on 11p, which is

known to contain multiple potential tumour suppresser genes [82], were simultaneously hypermethylated [81]. It was suggested that 11p region is a hot spot for CpG island methylation in neoplasia and that this DNA methylation change could be an important potential mechanism for inactivation of tumour suppresser genes [81].

At the same time Antiquera and Bird showed that multiple CpG islands, some associated with genes, were hypermethylated in immortalised human and murine cells [48]. It was postulated that as many as half of the CpG islands in the genome might be so altered in such cells [48,83]. A concept of methylation –associated gene inactivation (MAGI) was suggested [83].

Later loss of the expression, associated with hypermethylation of promoter CpG island was shown for retinoblastoma (Rb) gene in 10% of patients with sporadic form of retinoblastoma [84]. Several publications have documented *de novo* methylation of the CpG island for the cyclin dependent kinase inhibitor p16 in both cancer cell lines and primary tumours [85,86]. The aberrant hypermethylation correlated with the lack of p16 expression in these cells. Treatment of the cell lines with 5-aza-2'-deoxycytidine resulted in the demethylation of p16 promoter and reactivation of p16 expression [87].

In the case of colon cancer methylation of promoters of several genes was analysed. The particular site of interest in the case of colon carcinoma is the short arm of chromosome 11. Promoters of two genes located there, WT1 and calcitonin appeared to be hypermethylated in the majority of colon carcinomas (68-74%) [88]. Further, the promoter region of APC gene (Chromosome 5q21-q22) appears to be hypermethylated in more than half of patients with sporadic colorectal carcinoma, leading to the loss of the gene expression [89]. This hypermethylation was observed on the later stages of colon

carcinoma progression [89], suggesting the possible role of this effect not in the initiation, but in the progression of carcinoma. Expression of hMLH1 gene was shown to be suppressed by methylation [90]. In this case methylation of the hMLH1 promoter can cause the same effect as inactivating mutation leading to loss of functional gene [91]. Additional evidence for late role of methylation in the progression of colon carcinoma given by the results of Issa et al. [92] showing the increased activity of DNMT1 in carcinoma but not in adenoma.

One of the most striking system used for the analysis of the role of DNA methylation in carcinogenesis was an APC^{MIN} mouse model analysed by Laird et al. [93]. APC^{MIN} mouse carries a germline mutation in APC gene and develop hundreds of intestinal polyps. The inhibition of methyltransferase with 5-aza-2'-deoxycytidine in these mouse resulted in the reduction of polyp number from 113 to only 2. Although the mechanism by which demethylation reduces the polyp formation remains unclear, these results provide an important evidence for the involvement of methylation in carcinogenesis [93].

In summary increased methylation within the promoter regions of tumour suppressor genes can lead to the loss of their expression and thus contribute to tumour initiation or progression.

2 Objectives of this work

The present work was based on the observation that in the two types of colon tumours, nonmucinous and mucinous, the MUC2 gene was suppressed or strongly overexpressed, respectively, relative to the normal colon mucosa. This observation has led to the hypothesis that different regulation of the MUC2 gene in these tumours is a tumour-associated property, determining the mucinous versus nonmucinous pathways of colon carcinogenesis.

The objectives were:

To isolate and sequence the promoter of MUC2 gene.

To analyse the regulation of the gene in colorectal cancer cells *in vitro*.

The finding that methylation is the main regulatory mechanism prompted us to investigate in detail:

The effect of different methylation patterns on MUC2 expression *in vitro*.

The extent of MUC2 promoter methylation in the DNA of established colorectal cancer cell lines and of normal and carcinomatous human colon tissue.

The particular interest was focused again on the difference in methylation between the normal and the mucinous & nonmucinous carcinoma with the intention to prove or disprove the above-mentioned hypothesis.