The Royal Society

Control of gene expression in intestinal epithelial cells

Peter G. Traber

Department of Medicine, 100 Centrex, Hospital of the University of Pennsylvania, 34th and Spruce Streets, Philadelphia, PA 19104, USA (traberp@mail.med.upenn.edu)

Coordination of gene transcription is a critical regulatory step in orchestrating developmental, differentiation and adaptation processes in the mammalian intestinal epithelium. Insight into these mechanisms has been gained by the study of transcriptional regulation of the sucrase–isomaltase gene. An understanding of the regulatory network of nuclear proteins that direct transcriptional initiation of intestinal genes such as sucrase–isomaltase will provide insight into the mechanisms of normal development and differentiation as well as disease processes such as neoplasia.

Keywords: transcription; sucrase–isomaltase; transgenic mice; colon cancer; caudal-related homeobox genes; hepatocyte nuclear factor 1

1. INTRODUCTION

The intestinal epithelium is composed of multiple cell types that arise from a stem cell compartment located in crypts. The composition of cells in crypts and on the villus represents a complex equilibrium system that undergoes continual renewal while maintaining precise interrelationships. This epithelium develops via a series of transitions resulting from the interaction of visceral endoderm and mesoderm. The stratified, undifferentiated endoderm of the gut is converted into a simple columnar epithelium (endoderm–intestinal transition) relatively late in development at a time when the gross morphology of the alimentary tract is well defined. Also around this time the four different cellular phenotypes are established and nascent villi form. Thereafter crypts develop from the intervillus cells and villi lengthen until the adult phenotype is attained. In rodent intestine this process starts in late gestation and continues throughout the neonatal period, with completion in the third week of life around the time the pup is weaned from mother's milk. There are pronounced changes in gene expression at the time of this suckling–weaning transition that marks the beginning of the adult intestinal phenotype. An understanding of the mechanisms underlying these developmental changes will be important for deciphering the processes that maintain the precise cellular relationships and patterns of proliferation and differentiation in the adult epithelium.

During development, and in the adult epithelium, cellular phenotypes are defined by the expression of specific sets of genes in individual cells. The sets of genes expressed in intestinal epithelial cells are principally determined by transcriptional initiation, and the particular set of genes expressed in a single cell type has recently been referred to as the 'transcriptome'. Intestinal epithelial cell transcriptomes shift in well-orchestrated patterns during developmental, differentiation and adaptive processes of the intestinal mucosa. Thus, the molecular mechanisms that regulate transcription of cellular gene sets are the foundation for understanding developmental and differentiation events.

2. DEVELOPMENTAL REGULATION OF SUCRASE–ISOMALTASE GENE EXPRESSION

Current understanding of transcriptional regulatory mechanisms in the intestinal epithelium is based on intense scrutiny of a few genes expressed in restricted cellular and developmental patterns. Our studies on the regulatory mechanisms of the sucrase–isomaltase (*SI*) gene have underscored the complexity of understanding intestinal gene expression during development and differentiation. The *SI* gene is expressed during mouse intestinal development in a pattern that mirrors several critical developmental transitions (Markowitz *et al.* 1995; Tung *et al.* 1997). Late in gestation in the mouse, after the gross morphology of the intestinal tract has been established, the stratified visceral endoderm lining the immature gut undergoes a transformation to a simple columnar, polarized epithelium (Traber & Wu 1995). This endoderm–intestinal transition is a time when the various intestinal epithelial lineages are first detected and multiple intestinal genes are first expressed. *SI* mRNA is also first detected in the newly formed columnar cells, but the level of expression can only be detected using the most sensitive methods of mRNA measurement and SI protein is not detectable (Tung *et al.* 1997). *SI* expression remains at very low levels until the third postnatal week when the pups are in the process of being weaned from mother's milk to solid food. At this suckling–weaning transition there is a marked induction of *SI* mRNA expression to levels equivalent to those in adult epithelium (Markowitz *et al.* 1995; Tung *et al.* 1997). Multiple other genes are induced at this transition and there are changes in the morphology of the epithelial cells (Henning 1985, 1987). In the adult epithelium, *SI* mRNA and protein are expressed

predominantly in differentiated cells located on the villus. In the migration of epithelial cells from the base of crypts to villus tips, *SI* is first expressed in detectable levels in cells located in the upper crypt region and is expressed in high levels through the mid-villus with diminished levels in cells at the villus tips. Thus, complex patterns of *SI* expression are maintained in the continually renewing adult epithelium.

Transgenic mouse experiments show that the pattern of *SI* gene expression during development and in the mature epithelium is regulated largely at the level of gene transcription via multiple functional cis-acting DNA elements (Markowitz *et al.* 1993, 1995; Tung *et al.* 1997). Although these experiments indicate that the full complexity of the regulation of *SI* gene transcription remains to be described, important principles have been uncovered that will guide additional studies. Most importantly, the elements necessary to direct cell-specific expression in the intestinal epithelium appear to be embodied in a 201 nucleotide, evolutionarily conserved 5′-flanking region of the gene (Tung *et al.* 1997). Nucleotides −201 to +54 of the mouse *SI* gene directed expression of a human growth hormone reporter gene specifically to the intestinal epithelium. The predominant finding was high level expression of the transgene in villus-associated enterocytes appearing at the time of the suckling–weaning transition, similar to the pattern of expression of the endogenous gene. Thus, to a remarkable degree this short promoter was able to recapitulate the pattern of *SI* gene expression in enterocytes.

There were some notable differences between expression of the endogenous *SI* gene and the −201 *mSI* transgene. First, the level of expression of the transgene was variable between different transgenic lines and did not relate to the number of copies of the transgene incorporated in the genome. These data suggested that the expression of the *SI* gene promoter was affected by the site of chromosomal insertion and therefore dependent on the physical state of the surrounding DNA. In addition, there was ectopic expression of the transgene in a small percentage of enteroendocrine cells, with expression starting at the endoderm–intestinal transition in late foetal period. Experiments to examine the expression of a transgene construct with additional 5′-flanking DNA added information to the regulation of the gene in the intestinal mucosa (Markowitz *et al.* 1995). When nucleotides −8500 to +54 were linked to a reporter gene the pattern of expression was similar to the shorter construct until after the suckling–weaning transition. After induction of transgene expression in the enterocytes at the suckling–weaning transition, the transgene was expressed in Paneth cells, additional enteroendocrine cells and goblet cells, as well as enterocytes in crypts over the course of postnatal weeks 4 and 5 (Markowitz *et al.* 1995). These data suggested that all epithelial cells appear to have the transcriptional machinery necessary to express the *SI* gene (Markowitz *et al.* 1995). A transcriptional silencing element located outside the limits of the *SI* genomic DNA analysed might explain the discrepancy between expression of the transgene and the endogenous gene. An additional finding was that the expression of the long transgene in multiple transgenic lines was not dependent on the site of insertion and was related to the number of copies of the transgene inserted in the chromosomal DNA. These findings suggest that there may be DNA elements between nucleotides −8500 and −201 that are required to insulate the promoter from surrounding chromatin effects determined by the site of insertion of the transgene. These types of regulatory elements have been described for numerous genes and may be important standard components of gene structure (Wilson *et al.* 1990; Forrester *et al.* 1994; Festenstein *et al.* 1996).

3. MODULATION OF DEVELOPMENTAL GENE EXPRESSION BY CORTICOSTEROIDS

An important phenomenon in intestinal development is the ability of stress, via corticosteroids as the final mediator, to induce precocious maturation of the suckling rodent intestine (reviewed in Henning (1987)). Recently, evidence was presented that suggests this precocious maturation may be mediated through a separate pathway than that which directs the normal developmental process (Nanthakumar & Henning 1995). We found that neither −201, nor −8500 transgene constructs were responsive to induction by treatment of transgenic animals with dexamethasone (Tung *et al.* 1997). These data in transgenic mice provide molecular evidence that normal and precocious expression are indeed mediated via different pathways. Although the transgenic constructs included all of the elements required for developmental expression of *SI* in enterocytes, there appear to be DNA regulatory elements located outside nucleotides −8500 to +54 that are responsible for induction of *SI* gene transcription after treatment with dexamethasone. However, these data are unable to distinguish between a direct or indirect effect of the nuclear corticosteroid receptor on the *SI* gene.

4. MECHANISMS OF TRANSCRIPTIONAL INITIATION OF THE *SI* GENE

Transcriptional initiation is an intricate biochemical process that involves the interaction of core nuclear machinery (basal transcriptional apparatus) with cell-specific DNA binding proteins. Co-adaptor proteins serve to link the specific DNA binding proteins to portions of the basal apparatus. The coordinated assembly of a complex of many proteins at the transcriptional initiation site of a gene leads to modification of chromatin structure and initiation of RNA synthesis by RNA polymerase II. The biochemical interactions of these proteins are modulated by phosphorylation, a process that is connected to multiple cellular signalling pathways. Therefore, the regulatory networks of transcriptional proteins that control expression of intestinal epithelial genes are modulated by many processes that extend beyond simple expression of transcription factors.

DNA regulatory elements and associated DNA binding proteins have been carefully assessed within the evolutionarily conserved promoter region (Traber & Silberg 1996). There are at least three groups of transcriptional proteins involved in *SI* promoter transcription, including hepatocyte nuclear factor I (HNF1) (Wu *et al.* 1994), caudal-related homeodomain proteins (Cdx) (Traber *et al.* 1992; Suh *et al.* 1994), and nuclear proteins that interact with a GATA binding site which are as yet poorly characterized (Traber & Silberg 1996). HNF1α and HNF1β proteins

interact with two elements in the human *SI* gene promoter, *SIF2* and *SIF3* (Wu *et al*. 1994). HNF1α transactivates transcription of the *SI* promoter whereas HNF1β does not activate transcription, suggesting that relative levels of the two proteins may modulate transcriptional activation (Wu *et al*. 1994). Cdx1 and Cdx2 proteins, both of which are expressed in intestinal epithelial cells, bind to two closely apposed binding sites just upstream of the TATA box of the gene and activate transcription of the promoter (Suh *et al*. 1994).

5. TRANSCRIPTION FACTORS AND DEVELOPMENTAL PATTERNS OF *SI* GENE TRANSCRIPTION

Although a number of important transcription factors for *SI* gene transcription have been characterized, there is a lack of coordinated expression between these putative regulatory transcription factors and the *SI* gene. HNF1α and HNF1β mRNAs and proteins are expressed very early in the development of the gut tube, long before the endoderm–intestinal transition and expression is maintained into adulthood. Cdx1 and Cdx2 are expressed in the endoderm just before the endoderm–intestinal transition. This temporal relationship of expression and data demonstrating the effect of Cdx2 on intestinal gene expression and cellular morphogenesis (Suh & Traber 1996) suggests a vital role of caudal-related proteins in morphogenesis of the mucosa at this developmental transition. However, there is no clear change in the expression of Cdx mRNAs at the suckling–weaning transition, when there is marked induction of the *SI* gene. There are many possible reasons for these discrepancies, including: (i) differences in the relative levels or localization of proteins in individual cells; (ii) changes in the physical state of the chromatin that alters ability for transcription factors to interact with DNA; (iii) protein modifications affecting function such as phosphorylation; (iv) functional importance of other DNA binding factors that have yet to be discovered; or (v) coactivator proteins that may be required for linking factors bound to the promoter elements with the basal transcriptional apparatus. Each of these possibilities must be carefully examined to determine whether they may be important in the transcriptional patterns of the *SI* gene.

From the available data, we hypothesize that the temporal and spatial patterns of *SI* gene transcription in the intestinal epithelium is dependent on coordinated interactions of multiple DNA binding proteins that are linked to the basal transcriptional apparatus by coadaptor proteins. Modulation of these complex DNA–protein and protein–protein interactions may occur via phosphorylation of specific proteins. Protein modifications are likely linked to cellular signalling processes that are activated by extracellular cues such as matrix, growth factors or adjacent cells. Continued analysis of the *SI* gene promoter and its function during development will provide the framework to build a model of these complex interactions that regulate gene transcription.

6. TRANSCRIPTIONAL FACTORS AND INTESTINAL NEOPLASIA

There is evidence to suggest that an understanding of the transcriptional mechanisms of a single intestinal gene such as *SI* will have importance in understanding human disease such as colorectal neoplasia. Following the discovery that Cdx2 had a functional role in transcription of *SI*, it was found to have a profound effect on proliferation, morphogenesis and gene expression in intestinal epithelial cells (Suh *et al*. 1994). Subsequently, it has been shown that both Cdx2 and Cdx1 expression is markedly diminished in human colon adenomas and carcinomas (Ee *et al*. 1995; Mallo *et al*. 1997; Silberg *et al*. 1997). Finally, it was recently shown that mice harbouring one null allele for Cdx2 develop colorectal neoplasms early in life (Chawengsaksophak *et al*. 1997). Thus, diminished expression of Cdx2, and possibly Cdx1, may be involved in the pathogenesis of colorectal neoplasms. As the other pieces of the transcriptional puzzle of the *SI* gene are assembled there will be new information to generate hypotheses on the relationship of transcriptional mechanisms to intestinal disease.

REFERENCES

Chawengsaksophak, K., James, R., Hammond, V. E., Kontgen, F. & Beck, F. 1997 Homeosis and intestinal tumors in Cdx2 mutant mice. *Nature* **386**, 84–87.

Ee, H. C., Erler, T., Bhathal, P. S., Young, G. P. & James, R. J. 1995 Cdx-2 Homeodomain protein expression in human and rat colorectal adenoma and carcinoma. *Am. J. Path.* **147**, 586–592.

Festenstein, R., Tolaini, M., Corbella, P., Mamalaki, C., Parrington, J., Fox, M., Miliou, A., Jones, M. & Kioussis, D. 1996 Locus control region function and heterochromatin-induced position effect variegation. *Science* **271**, 1123–1125.

Forrester, W. C., VanGenderen, C., Jenuwein, T. & Grosschedl, R. 1994 Dependence of enhancer-mediated transcription of the immunoglobulin μ gene on nuclear matrix attachment regions. *Science* **265**, 1221–1225.

Henning, S. J. 1985 Ontogeny of enzymes in the small intestine. *A. Rev. Physiol.* **47**, 231–245.

Henning, S. J. 1987 Functional development of the gastrointestinal tract. In *Physiology of the gastrointestinal tract* (ed. L. R. Johnson), pp. 285–300. New York: Raven.

Mallo, G. V., Rechreche, H., Frigerio, J.-M., Rocha, D., Zweibaum, A., Lacasa, M., Jordan, B. R., Busetti, N. J., Dagorn, J.-C. & Iovanna, J. L. 1997 Molecular cloning, sequencing, and expression of the mRNA encoding human CDX1 and CDX2 homeobox, down-regulation of CDX1 and CDX2 mRNA expression during colorectal carcinogenesis. *Int. J. Cancer* **74**, 35–44.

Markowitz, A. J., Wu, G. D., Bader, A., Cui, Z., Chen, L. & Traber, P. G. 1995 Regulation of lineage-specific transcription of the sucrase–isomaltase gene in transgenic mice and cell lines. *Am. J. Physiol.* **269**, G925–G939.

Markowitz, A. J., Wu, G. D., Birkenmeier, E. H. & Traber, P. G. 1993 The human sucrase–isomaltase gene directs complex patterns of gene expression in transgenic mice. *Am. J. Physiol.* **265**, G526–G539.

Nanthakumar, N. N. & Henning, S. J. 1995 Distinguishing normal and glucocorticoid-induced maturation of intestine using bromodeoxyuridine. *Am. J. Physiol.* **268**, G139–G145.

Silberg, D. G., Furth, E. E., Taylor, J. K., Schuck, T., Chiou, T. & Traber, P. G. 1997 Cdx1 protein expression in normal, metaplastic and neoplastic human alimentary tract epithelium. *Gastroenterology* **113**, 478–486.

Suh, E.-R. & Traber, P. G. 1996 An intestine-specific homeobox gene regulates proliferation and differentiation. *Molec. Cell. Biol.* **16**, 619–625.

Suh, E.-R., Chen, L., Taylor, J. & Traber, P. G. 1994 A homeodomain protein related to caudal regulates intestine-specific gene transcription. *Molec. Cell. Biol.* **14**, 7340–7351.

Traber, P. G. & Silberg, D. G. 1996 Intestine-specific gene transcription. *A. Rev. Physiol.* **58**, 275–297.

Traber, P. G. & Wu, G. D. 1995 Intestinal development and differentiation. In *Gastrointestinal cancers: biology, diagnosis, and therapy* (ed. A. K. Rustgi), pp. 21–43. Philadelphia, PA: Lippincott–Raven.

Traber, P. G., Wu, G. D. & Wang, W. 1992 Novel DNA-binding proteins regulate intestine-specific transcription of the sucrase–isomaltase gene. *Molec. Cell. Biol.* **12**, 3614–3627.

Tung, J., Markowitz, A. J., Silberg, D. G. & Traber, P. G. 1997 Developmental expression of SI in transgenic mice is regulated by an evolutionarily conserved promoter. *Am. J. Physiol.* **273**, G83–G92.

Wilson, C., Bellen, H. J. & Gehring, W. J. 1990 Position effects on eukaryotic gene expression. *A. Rev. Cell Biol.* **6**, 679–714.

Wu, G. D., Chen, L., Forslund, K. & Traber, P. G. 1994 Hepatocyte nuclear factor 1α (HNF-1α) and HNF-1β regulate transcription via two elements in an intestine-specific promoter. *J. Biol. Chem.* **269**, 17 080–17 085.