Coordinate Regulation of the Gel-forming Mucin Genes at Chromosome 11p15.5*

Received for publication, November 21, 2012, and in revised form, December 24, 2012. Published, JBC Papers in Press, January 9, 2013, DOI 10.1074/jbc.M112.437400

Nehal Gosalia1, Shih-Hsing Leir1, and Ann Harris1,6

From the 1Human Molecular Genetics Program, Children’s Memorial Research Center and Department of Pediatrics and 6Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, Illinois 60614

Background: The gel-forming mucin genes at chr11p15.5 encode major components of the epithelial mucus layer.

Results: CTCF occupancy correlates with mucin gene expression and active loci form three-dimensional looped structures.

Conclusion: CTCF has a pivotal role in coordinating gene expression and response to lipopolysaccharide.

Significance: The study provides novel therapeutic prospects for epithelial diseases associated with mucin hypersecretion.

Four of the genes that encode gel-forming mucins, which are major components of the mucus layer protecting many epithelial surfaces, are clustered at chromosome 11p15.5 and show both cell- and tissue-specific expression patterns. We aimed to determine whether the individual genes were coordinately regulated by mechanisms involving higher order chromatin structure. CCCTC-binding factor (CTCF) sites were predicted in silico and CTCF occupancy then evaluated by chromatin immunoprecipitation. CTCF was found at many sites across the gene cluster, and its binding was correlated with mucin gene expression. Next, siRNA-mediated depletion of CTCF was shown to increase MUC2 expression in A549 lung carcinoma cells and both MUC6 and MUC5AC expression in LS180 colon carcinoma cells. These changes correlated with loss of CTCF binding at multiple sites, although others retained occupancy. In cells actively expressing the mucins, the gene cluster was shown by chromosome conformation capture to form looped three-dimensional structures with direct interactions between the MUC2 promoter region, regions 30 kb 5’ to it, close to the MUC6 promoter and others near the 3’ end of MUC5AC, >170 kb away. Finally, to demonstrate the importance of CTCF binding to mucin gene expression, Calu-3 lung carcinoma cells were exposed to lipopolysaccharide (LPS). LPS increased the expression of MUC2 and MUC5AC and reduced MUC5B. CTCF occupancy was concurrently depleted at specific binding sites close to these genes. These data suggest that CTCF binding and cell type-specific long-range interactions across the 11p15.5 gene cluster are critical mechanisms for coordinating gel-forming mucin gene expression.

Recent advances in the methodology for analysis of transcriptional regulatory mechanisms genome-wide are revealing many novel functions for non-coding sequences (1). These elements are generally cis-acting but may be located at a considerable distance (>100 kb) from the genes they control. The new insights into the control of expression of individual genes are extensive; however, the opportunity to reveal regulatory mechanisms for gene clusters and gene networks across the genome is particularly exciting. Many functionally linked gene clusters such as the homeobox, β-globin, and gel-forming mucin loci are thought to have evolved by duplication of an ancestral gene followed by evolution of genes with both common and divergent properties (2, 3). The mechanisms that subsequently coordinate the whole gene cluster in addition to the individual loci within it are of considerable interest. We sought to determine the regulatory mechanisms controlling gene expression within the gel-forming mucin gene cluster located at chromosome 11p15.5 because synchronized expression of these genes is likely critical for normal function of epithelia lining many organs. Mucins are high molecular weight glycoproteins that contain tandem repeat regions enriched for serine, threonine, and proline, which are modified by O-glycosylation. These glycoproteins may be membrane-bound or secreted (gel-forming), and together, both types of mucin generate the characteristic mucus layer that protects epithelial surfaces in the intestine, pancreas, kidney, lungs, and many other organs (4, 5). Mucins also have additional critical functions in vivo and their misregulation is a characteristic feature of many disorders, including chronic obstructive pulmonary disease, cystic fibrosis, asthma, gastrointestinal inflammatory diseases, and lung, colon, and pancreatic carcinomas (6–10). The four gel-forming mucin genes at chromosome 11p15.5 include in order (telomeric to centromeric) MUC6, MUC2, MUC5AC, and MUC5B and encompass a ~400-kb region (11). The expression of the four mucin genes is cell type- and tissue-specific: MUC6 is one of the two major gastric mucins (12) and is restricted to neck mucous cells (13). MUC2 is highly expressed in goblet cells within the intestinal epithelium and at lower levels in the epithelium lining conducting airways (8, 14); MUC5AC is expressed in surface goblet cells in the airway and the stomach (8, 15, 16) where its pattern of expression does not overlap with MUC6; and MUC5B expression is largely localized to submucosal gland cells within the respiratory tract (8, 16, 17).

Extensive analysis of the transcriptional regulation of individual genes encoding the gel-forming mucins has already been pursued and demonstrated that many of the gene promoters...
Higher Order Chromatin Structure of the Gel-forming Mucin Gene

are activated by similar mechanisms. These include inflammatory mediators such as cytokines (interleukin 1β (IL-1β), IL-4, IL-6, IL-9, IL-12, IL-13 and tumor necrosis factor α (TNFα)); growth factors (epidermal growth factor (EGF), transforming growth factor α (TGF-α)) and bacterial LPS (reviewed in Ref. 10). Moreover, the four gene promoters contain functional binding sites for many general transcription factors, including Sp1, Sp3, CREB, AP-1, NF-κB, and c-Myc (reviewed in Refs. 10, 18) as would be expected from their mechanisms of activation. Recruitment of diverse transcription factors to these mucin genes results in their epigenetic modification by mechanisms, including DNA methylation, histone methylation, and histone acetylation/deacetylation (19, 20). Despite this wealth of information on the regulation of expression of MUC6, MUC2, MUC5AC, and MUC5B, the mechanisms that coordinate the expression of the whole gene cluster have not yet been examined. However, these more global regulatory mechanisms are critically important in ensuring the appropriate composition of the mucus gel in specific epithelial environments and that a well orchestrated response is generated to epithelial insults. Because individual mucus glycoproteins have quite different biochemical and biophysical properties, it is clear that misexpression will have functional consequences.

We tested the hypothesis that the 11p15.5 mucin gene cluster is coordinately regulated by modifications in higher order chromatin structure that enhance or restrict access to individual gene promoters within the cluster. Moreover, that the three-dimensional organization of this genomic region modulates the interaction of cis-acting enhancers or repressors with the individual gene promoters. Long range interactions are known to regulate other multigene clusters, including the CTCF, CCCTC-binding factor (CTCF), interaction with insulator elements and prevents the spread of heterochromatin (27). CTCF is an 82-kDa protein that binds to protein complexes; it also isolates distinct chromatin domains by associating with boundary elements and prevents the spread of heterochromatin (27). CTCF is an 82-kDa protein that binds to protein complexes; it also isolates distinct chromatin domains by associating with boundary elements and prevents the spread of heterochromatin (27). CTCF is an 82-kDa protein that binds to protein complexes; it also isolates distinct chromatin domains by associating with boundary elements and prevents the spread of heterochromatin (27). CTCF is an 82-kDa protein that binds to protein complexes; it also isolates distinct chromatin domains by associating with boundary elements and prevents the spread of heterochromatin (27). CTCF is an 82-kDa protein that binds to protein complexes; it also isolates distinct chromatin domains by associating with boundary elements and prevents the spread of heterochromatin (27). CTCF is an 82-kDa protein that binds to protein complexes; it also isolates distinct chromatin domains by associating with boundary elements and prevents the spread of heterochromatin (27).
interacting region primers used are listed in supplemental Table 1.

Lipopolysaccharide Treatment—Calu-3 cells were serum-starved from 24 h and treated either with PBS or with 200 ng/ml Pseudomonas aeruginosa LPS in PBS from Sigma (L9134).

RESULTS

Differential Mucin Gene Expression Patterns in Colon, Lung, and Pancreatic Adenocarcinoma Cell Lines—The chr.11p15.5 gel-forming mucins have distinct expression patterns that vary by cell and tissue type. In normal cells, only one gel-forming mucin is usually expressed (13); however, in cancer cell lines, multiple mucin genes may be expressed, thus providing a more informative model in which to examine coordinate gene regulation across the cluster. Although mucin gene expression data exist for many cell types, these are rarely quantitative. Hence, we first generated robust qRT-PCR assays for each of the four gel-forming mucin genes. To facilitate comparison of individual mucin genes within a cell type, we measured mucin gene expression using absolute quantification (Fig. 1). For each cell type, the individual expression values, normalized to 18 S rRNA, were set relative to LS180, a colon adenocarcinoma cell line that expresses all four genes. Next, picomolar amounts of the cDNA for each gel-forming mucin gene in LS180 were calculated. These absolute amounts were multiplied by the relative expression values of each mucin in the other cell lines. For example, if cell line A expressed 2-fold more MUC6 than LS180, using the picomolar amount for MUC6 in LS180 (~0.1), line A would show an expression value of ~0.2. The relative and absolute expression data sets, both show tissue- and cell type-specific differences in mucin gene expression. To illustrate, the lung carcinoma cell lines A549 and Calu-3 both express MUC5AC, but only A549 expresses MUC5B. Moreover, although LS180 and LS174T are derived from the same colon carcinoma cell line, LS174T expresses 10-fold more MUC2 (Fig. 1). These distinct gel-forming mucin gene expression profiles and the expression of several mucins by a single cancer cell suggest that there may be a mechanism to facilitate expression of certain genes and the repression of others.

The Mucin Gene Cluster Is Demarcated by Multiple CTCF Binding Sites—The mucin gene cluster at chromosome 11p15.5 encompasses ~400 kb and contains four genes in the order MUC6, MUC2, MUC5AC, and MUC5B (Fig. 2A), with MUC6 telomeric and MUC5B centromeric. MUC6 is transcribed on the reverse strand, and MUC2, MUC5AC, and MUC5B are transcribed on the forward strand. There is a known gap in the HG19 assembly of the human genome sequence within the tandem repeat region of the MUC5AC gene (marked with dashed line on Fig. 2A). However, this does not impact interpretation of our data because the cDNA of MUC5AC has been cloned (44), enabling accurate expression studies and primer sets used in q3C lie outside this region.

CTCF is known to regulate many multigene clusters to facilitate appropriate expression of each gene through modifying higher order chromatin structure and/or by binding to enhancer-blocking insulator elements. Thus, we used chromatin immunoprecipitation (ChIP) with an antibody specific for CTCF to examine its occupancy at predicted binding sites across the cluster in relevant cell types. Thirteen CTCF binding sites of 26 seen in ENCODE data (1) (derived from 47 cell lines) were analyzed by qRT-PCR (Fig. 2B, sites I–XIII). The choice of sites for analysis was made based on their location (all are intronic or intergenic), and CTCF occupancy was shown in relevant cell types. Among these sites several, for example sites IV and IX, show ubiquitous occupancy of CTCF in all cell types, although the levels of enrichment are variable. At other sites such as VII, XI, and XII, CTCF occupancy differs greatly between cell types, ranging from very low enrichment to the highest level of occupancy of all sites in a cell line (compare site XI in Calu-3 and A549). Interestingly, sites V, VI, and VIII, which are predicted to bind CTCF based on ENCODE data, show no enrichment despite adjacent sites showing high occupancy (e.g. site IV). The CTCF binding profile shows some correlation with mucin gene expression patterns, irrespective of the tissue origin of the cell line. For example, in Calu-3 (airway) and HT-29 (colon) cell lines, which express only MUC5AC (Fig. 1), the overall CTCF binding profile is very similar with the exception of site XI, at the 5’ end of MUC5B. In contrast, if both MUC5AC and MUC5B are expressed, as in A549 (airway) and Capan-1 (pancreatic) lines, CTCF occupancy at sites near the 5’ end (XI) and 3’ (XII) to MUC5B decreases (Fig. 2B). LS180 and LS174T, which are derivatives of the same colon carcinoma cell line, express all four gel-forming mucins. However, MUC2 expression levels are higher in LS174T (Fig. 1), and this apparently correlates with reduced CTCF occupancy at sites between MUC6 and MUC2 (sites III, IV) in comparison with LS180. The pancreatic cells express either very high levels of MUC2 (BxPC-3) or MUC6 (Capan-1) but not both. In these lines, there
is overlap between sites of CTCF occupancy even though the gene expression profile is different. However, overall occupancy of CTCF across the gene cluster is higher in BxPC-3, which expresses very low levels of all the other genes, in contrast to Capan-1, which expresses several genes. Having established the CTCF occupancy profile of each cell line in relation to its expression of the gel-forming mucin genes, we next examined the effect of altering CTCF binding profiles on gene expression patterns.

Changes in CTCF Occupancy Alter Gel-forming Mucin Gene Expression—To determine whether CTCF occupancy at sites within the chr11p15.5 mucin gene cluster impacts gene expression, we used specific siRNAs to deplete CTCF. We first utilized A549 lung carcinoma cells, which express MUC5AC and MUC5B (Fig. 1). Efficient (~90%) knockdown of CTCF resulted in an ~3-fold increase in MUC2 expression (p < 0.05) in comparison with cells treated with a non-targeting siRNA control (Fig. 3A). Concurrent slight increases in MUC5AC and decreases in MUC5B expression after CTCF depletion were not statistically significant, and MUC6 levels did not change. To determine whether these alterations in mucin gene expression correlated with changes in CTCF occupancy at any sites, we performed ChIP for CTCF in negative control siRNA and CTCF siRNA-treated cells (Fig. 3B). Multiple CTCF-binding sites across the gene cluster showed selective depletion of CTCF occupancy after siRNA treatment. Of particular interest was loss of enrichment at sites VI and VII within the MUC2 gene, consistent with a dramatic increase in its expression, and at site IX, which is located within MUC5AC and close to the 5' end of MUC5B because both genes are highly expressed in A549 cells. In contrast, CTCF knockdown had only a slight impact on occupancy at site IV, which lies between MUC6 and
MUC2. This observation may underlie the apparent lack of transcriptional response of the MUC6 gene to CTCF depletion.

Next, we evaluated the effect of CTCF knockdown in LS180 colon carcinoma cells, which have a different mucin gene expression pattern from A549 and express all four gel-forming mucin genes (Fig. 1). siRNA-mediated CTCF depletion resulted in a significant increase in both MUC6 (p < 0.01) and MUC5AC (p < 0.001) expression relative to cells treated with a non-targeting siRNA (Fig. 3C). To identify the changes in CTCF occupancy after its depletion, we again performed ChIP for CTCF in negative control and CTCF siRNA-treated cells. Most sites showed loss of CTCF enrichment though the relative levels of depletion vary by site (Fig. 3D). None of the changes showed close correlation with expression of MUC6 and MUC5AC with the exception of site VII, perhaps because both genes are already expressed in this line. Because CTCF knockdown in both A549 and LS180 cells increased mucin gene expression, this factor may be establishing a repressive chromatin environment at the gene cluster.

Long Range Interactions at the Gel-forming Mucin Gene Cluster—We next used q3C to test the hypothesis that some aspects of coordinate regulation of the mucin genes at chr11p15.5 are mediated by the long range interactions of cis-acting elements across the cluster. We first measured interactions between the MUC2 promoter and other regions in the gene cluster. In skin fibroblasts, which do not express any gel-forming mucins, there are no interactions across the gene cluster, except for the expected interaction of regions < 10 kb away from the MUC2 promoter bait (Fig. 4). In contrast, in LS174T and LS180 cells, which express high levels of all four genes, higher interactions were observed between the MUC2 promoter and elements close to the MUC6 5’ end (marked by black arrows in Fig. 4 at chr11:1,031,000–1,038,000). In LS174T cells, which express substantially higher levels of MUC2 (10-fold) than LS180 cells, the interaction frequencies across the region appeared greater. In addition, in LS180, a higher interaction frequency was seen between the MUC2 promoter and the 3’ end of the AP2A2 gene (at chr 11:1,005,000, gray arrow). In Capan-1 cells, which express abundant MUC6 but no MUC2, the interaction frequency between the MUC2 promoter and the MUC6 5’ end (chr11:1,038,000) was low, though slightly higher with the 3’ end of the AP2A2 gene (chr11:1,005,000). The interactions between the MUC6 and MUC2 promoters in cell lines that express both genes were confirmed with a 3C bait at the MUC6 gene promoter (supplemental Fig. 2). Interactions between the MUC2 promoter and the 3’ end of the mucin gene cluster were generally very low, with the exception of one region close to the 3’ end of the MUC5AC gene (chr11:1,218,000, light gray arrow) in Capan-1 cells. Capan-1 express much higher levels of MUC5AC and particularly MUC5B than LS180 or LS174T, and this may be relevant to the three-dimensional structure of the mucin gene cluster in these cells. Consistent with this hypothesis, A549 airway epithelial cells, which express high levels of MUC5AC and MUC5B show interactions between the MUC2 promoter and regions within the MUC5AC and MUC5B genes (Fig. 4, chr11:1,218,000 and 1,242,000, light gray arrows).

Lipopolysaccharide Exposure Decreases CTCF Occupancy across the Cluster in Association with Activation of Mucin Gene Expression—Among the many known in vivo activators of gel-forming mucin gene expression are bacterial exoproducts such as LPS, a component of the outer membrane of Gram-negative bacteria. For each mucin gene data were normalized by setting negative (Neg) control siRNA expression to 1. B and D, selective loss of CTCF occupancy after knockdown (B, A549; D, LS180). ChIP for CTCF was carried out on cells transfected with a non-targeting (control) siRNA or an siRNA targeting CTCF. Data are presented as fold enrichment over rabbit IgG isotype control. ChIP was done at least twice, and representative data (consistent profiles in all replicates) are shown.

**FIGURE 3.** siRNA-mediated depletion of CTCF augments specific mucin gene expression and causes differential loss of CTCF occupancy. A and C, changes in mucin gene expression after CTCF knockdown (KD; A, A549; C, LS180) were determined by RT-qPCR in negative control siRNA and CTCF siRNA treated cells. For each mucin gene data were normalized by setting negative (Neg) control siRNA expression to 1. B and D, selective loss of CTCF occupancy after knockdown (B, A549; D, LS180). ChIP for CTCF was carried out on cells transfected with a non-targeting (control) siRNA or an siRNA targeting CTCF. Data are presented as fold enrichment over rabbit IgG isotype control. ChIP was done at least twice, and representative data (consistent profiles in all replicates) are shown.
bacteria. To elucidate the biological relevance of CTCF binding at the gel-forming mucin cluster, we exposed Calu-3 cells to 200 ng/ml LPS from *P. aeruginosa*. This led to a significant increase in MUC2 and MUC5AC expression as expected (50–52), a significant (~25%) decrease in MUC5B expression (Fig. 5A), and no change in MUC6 expression (data not shown). Next, we measured CTCF occupancy by ChIP across the mucin gene cluster in untreated and LPS-treated cells. The results showed that LPS activation was accompanied by a partial loss of CTCF binding at several sites across the cluster, including site IV, which is located between MUC6 and MUC2 and sites XI and XII, which are located at the 5' and 3' end of MUC5B, respectively (Fig. 5B). These data suggest that loss of CTCF binding at the gel-forming mucin gene cluster is an integral part of the activation of these genes in response to LPS treatment.

**DISCUSSION**

Coordinate regulation of multigene clusters such as the gel-forming mucin gene complex on chromosome 11p15.5 often involves organization of the chromatin into higher order structures, which may be mediated by CTCF. Here, we examined the mechanistic role of CTCF in coordinating the complex pattern of regulation of the gel-forming mucin genes and mediating the long range interactions across the cluster. First, we profiled the expression of the four genes, MUC6, MUC2, MUC5AC, and MUC5B in colon, lung, and pancreatic cancer cells, for the first time using absolute quantification of transcripts. Next, we determined CTCF binding in these cell types by ChIP and found that in cells where the mucin gene cluster was active, CTCF bound at multiple sites. In contrast in fibroblasts, where the mucin genes are not expressed, CTCF binding is only seen at ubiquitous sites, and the cluster is marked by repressive histone modifications such as H3K9Me3 and H3K27Me3 (ENCODE data in supplemental Fig. 1). Interestingly, the predicted CTCF sites in the genes adjacent to the mucins, AP2A2 5' and Toll interacting protein (TOLLIP) 3', show relatively low occupancy in the cancer cell lines that we examined (Fig. 2B).
Two CTCF binding sites (IV and IX) showed occupancy in all mucin-expressing cell lines, irrespective of gene expression patterns, and these likely correspond to ubiquitous sites observed across the genome. However, several sites showed selective CTCF occupancy that correlated, at least in part, with gene expression profiles (Table 1). For example, Calu-3 and HT-29 express only MUC5AC at high levels (Fig. 1), and CTCF occupancy in both cell types is similar (Fig. 2B) with the exception of site XI. In comparison in A549 and Capan-1 cells, which both express MUC5B and MUC5AC, lower CTCF occupancy is seen at sites XI and XII, which are located near the 5’ (XI) and 3’ (XII) end of MUC5B (Fig. 2B). LS180 and LS174T express all four mucin genes, although at different abundance, and our absolute expression data showed that LS174T has 10-fold more MUC2 transcript compared with LS180 (Fig. 1). Concurrently, CTCF occupancy shown by ChIP at sites III and IV, which are located between MUC6 and MUC2, is decreased in LS174T in comparison with LS180 (Fig. 2B).

siRNA-mediated depletion of CTCF in A549 cells increased MUC2 expression ~3-fold (Fig. 3A) and in LS180 increased both MUC6 and MUC5AC expression (Fig. 3C). ChIP for CTCF subsequent to its knockdown showed loss of enrichment at most sites in both cell lines, although the effect was generally less at one ubiquitous CTCF site (IV) than at the cell-selective CTCF binding sites (Fig. 3, B and D). These data show that CTCF is a negative regulator of gel-forming mucin gene expression. It is likely based on data from other loci (29, 53, 54) that CTCF suppresses gene expression by recruiting a repressive complex such as the polycomb group repressor complex 2 or SIN3A and associated histone deacetylases. Consistent with these observations, previous reports showed an increase in MUC2 expression following histone deacetylase 2 depletion (19) and that expression of MUC2 was apparently dependent on H3K9 and H3K27 acetylation in the 5’-flanking region (55). However, our data showing that CTCF depletion increases expression of MUC2, MUC5AC, and MUC6 in different cell lines are in contrast with observations that only MUC2 and MUC5B are subject to epigenetic regulation, with MUC5AC being rarely influenced and MUC6 not at all (19).

To elucidate the biological relevance of CTCF binding at the 11p15.5 mucin gene cluster, LPS from P. aeruginosa was used to modulate mucin gene expression. LPS treatment significantly increased MUC2 and MUC5AC expression (Fig. 5A), and this correlated with and likely resulted from decreased CTCF binding at several sites across the cluster (Fig. 5B). Moreover, LPS treatment significantly reduced MUC5B expression (~25%), suggesting a coordinated regulatory mechanism for the gene cluster, whereby up-regulation of one or more gel-forming mucins is accompanied by down-regulation of others. The mechanism whereby MUC5B expression is decreased may involve the establishment of repressive histone modifications, consistent with previous observations on the regulation of MUC5B (19). These results are consistent with our data on siRNA-mediated depletion of CTCF, which show that CTCF is a repressor of mucin gene expression (Fig. 3). They also suggest that one mechanism of LPS-induction of gel-forming mucin gene expression is the inhibition of CTCF binding at certain sites across the cluster. This inhibition may also involve direct or indirect interactions with NF-κB at the mucin gene cluster. The impact of LPS activation on individual mucin gene promoters was studied previously (56, 57). P. aeruginosa LPS interacts with Toll-like receptor 4, which activates the Src/Ras/MAPK/pp90rsk pathway resulting in nuclear localization of NF-κB and its binding to elements close to the MUC2 and MUC5AC promoters (reviewed in Ref. 10). Our results suggest that LPS may have a more global role across the gene cluster. These data also suggest that CTCF may play an important role in other inflammatory pathways that up-regulate mucin gene expression but are activated by agents other than LPS.

We investigated long range interactions across the gel-forming mucin gene cluster by q3C and found higher order structures across the region. Using the MUC2 promoter as the bait region, we observed interactions with the 5’ end of MUC6 and the 3’ end of AP2A2 (Fig. 4). The MUC2 promoter interaction with the 5’ end of MUC6 may depend on both genes being expressed, as seen in LS180 and LS174T. If MUC2 is silent (Capan-1 and A549), its promoter does not interact with the 5’ end of MUC6 but still interacts with the 3’ end of AP2A2 (Fig. 4). Another long range interaction that seems dependent on gene activity is between the MUC2 promoter and regions close to the 3’ end of MUC5AC and the 5’ end of MUC5B in Capan-1 and A549. These long range interactions may be critical for facilitating expression of the gel-forming mucin genes, creating chromatin loops to prevent inappropriate cross-regulation between adjacent genes such as MUC6 and MUC2 and at the same time providing feedback to communicate gene expression patterns to other loci in the cluster. The characterization of an
extra level of transcriptional regulation across the gel-forming mucin gene cluster, in addition to the control of individual genes by their respective promoters, may be relevant to human disease. Many diseases of the airway epithelium such as chronic obstructive pulmonary disease, asthma, and cystic fibrosis, gastrointestinal inflammatory diseases, and carcinomas are associated with gel-forming mucin hypersecretion. The identification of CTCF as a critical mediator in controlling gene expression at this cluster may provide novel therapeutic opportunities because the protein is known to recruit repressive complexes that are suitable for pharmacological manipulation.

Acknowledgments—We thank Drs. Gunnar Hansson (MUC2), Thécla Lesuffleur (MUC5AC), and Marie-Pierre Buisine and Isabelle Van Seuningen (MUC5B) for the mucin cDNA clones.

REFERENCES

1. Rosenkranz, K. R., Dresser, T. R., Pheasant, M., Barber, G. P., Meyer, L. R., Pohl, A., Raney, B. J., Wang, T., Hinrichs, A. S., Zweig, A. S., Fujita, P. A., Learned, K., Rhead, B., Smith, K. E., Kuhn, R. M., Karolchik, D., Haussler, D., and Kent, W. J. (2010) ENCODE whole-genome data in the UCSC Genome Browser. Nucleic Acids Res. 38, D620–625

2. Sproul, D., Gilbert, N., and Bickmore, W. A. (2005) The role of chromatin structure in regulating the expression of clustered genes. Nat. Rev. Genet. 6, 775–781

3. Dessey, J. L., Aubert, J. P., Pochert, N., and Laine, A. (2000) Evolution of the Large Secreted Gel-Forming Mucins. Mol. Biol. Evol. 17, 1175–1184

4. Thornton, D. J., Rousseau, K., and McGuckin, M. A. (2008) Structure and Function of the Polymeric Mucins in Airways Mucus. Annu. Rev. Physiol. 70, 459–486

5. Hollingsworth, M. A., and Swanson, B. J. (2004) Mucins in cancer: protection and control of the cell surface. Nat. Rev. Cancer 4, 45–60

6. Tam, A., Wadsworth, S., Dorscheid, D., Man, S. F., and Sin, D. D. (2011) Mucin production during prenatal and postnatal murine lung development. Am. J. Respir. Cell Mol. Biol. 44, 755–760

7. Reid, C. J., Gould, S., and Harris, A. (1997) Developmental Expression of Mucin Genes in the Human Respiratory Tract. Annu. J. Respir. Cell Mol. Biol. 17, 592–598

8. Thai, P., Loukoianov, A., Wachi, S., and Wu, R. (2008) Regulation of Airway Mucin Gene Expression. Annu. Rev. Physiol. 70, 405–429

9. Vincent, A., Perrais, M., Desseyn, J. L., Aubert, J. P., Pigny, P., and Van Seuningen, I. (2007) Epigenetic regulation (DNA methylation, histone modifications) of the 11p15 mucin genes (MUC2, MUC5AC, MUC5B, MUC6) in epithelial cancer cells. Oncogene 26, 6566–6576

10. Yamada, N., Kitamoto, S., Yokoyama, S., Hamada, T., Goto, M., Tsutsu-mida, H., Higashi, M., and Yonezawa, S. (2011) Epigenetic regulation of mucin genes in human cancers. Clin. Epigenetics 2, 85–96

11. Spilianakis, C. G., and Flavell, R. A. (2004) Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. Nat. Immunol. 5, 1017–1027

12. Vennimmen, D., De Gobbi, M., Sloane-Stanley, J. A., Wood, W. G., and Higgs, D. R. (2007) Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. EMBO J. 26, 2041–2051

13. Majumder, P., Gomez, J. A., Chadwick, B. P., and Boss, J. M. (2008) The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. J. Exp. Med. 205, 785–798

14. Majumder, P., and Boss, J. M. (2010) CTCF controls expression and chromatin architecture of the human major histocompatibility complex class II locus. Mol. Cell Biol. 30, 4211–4223

15. Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F., and de Laat, W. (2002) Looping and interaction between hypersensitive sites in the active β-globin locus. Mol. Cell 10, 1453–1465

16. Hsu, C., Dale, R., and Dean, A. (2010) Cell type specificity of chromatin organization mediated by CTCF and cohesin. Proc. Natl. Acad. Sci. U.S.A. 107, 3651–3656

17. Krivega, I., and Dean, A. (2012) Enhancer and promoter interactions-long distance calls. Curr. Opin. Genet. Dev. 22, 79–85

18. Lobanenkov, V. V., Nicolas, R. H., Adler, V. V., Paterson, H., Klenova, E. M., Polotskaja, A. V., and Goodwin, G. H. (1990) A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5’-flanking sequence of the chicken c-myc gene. Oncogene 5, 1743–1753

19. Filippova, G. N., Fagerlie, S., Klenova, E. M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P. E., Collins, S. J., and Lobanenkov, V. V. (1996) An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Mol. Cell Biol. 16, 2802–2813

20. Ohlsson, R., Renkawitz, R., and Lobanenkov, V. (2001) CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends in Genetics 17, 520–527

21. Bell, A. C., West, A. G., and Felsenfeld, G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell 98, 387–396

22. Phillips, J. E., and Corces, V. G. (2009) CTCF: master weaver of the genome. Cell 137, 1194–1211

23. Boyle, A. P., Song, L., Lee, B. K., London, D., Keefe, D., Birney, E., Iyer, V. R., Crawford, G. E., and Furey, T. S. (2011) High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. Genome Res. 21, 456–464

24. Kim, T. H., Abdullaev, Z. K., Smith, A. D., Ching, K. A., Loukonin, D. I., Green, R. D., Zhang, M. Q., Lobanenkov, V. V., and Ren, B. (2007) Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 128, 1231–1245
Higher Order Chromatin Structure of the Gel-forming Mucin Gene

36. Cuddapah, S., Jothi, R., Schones, D. E., Roh, T. Y., Cui, K., and Zhao, K. (2009) Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome Res. 19, 24–32

37. Tom, B. H., Rutzky, L. P., Jakstys, M. M., Oyasu, R., Kaye, C. I., and Kahan, B. D. (1976) Human colonic adenocarcinoma cells. I. Establishment and description of a new line. In. Vitro. 12, 180–191

38. Fogh, J., and Trempe, G. (1975) New human tumor cell lines in Human Tumor Cells in Vitro. (Fogh, J., Ed.) pp. 115–141, Plenum Publishing Corp., New York

39. Lieber, M., Smith, B., Szakal, A., Nelson-Rees, W., Todaro, G. (1976) A clone isolated from HT-29 mucus-secreting cells. The 3

40. Lieber, M., Smith, B., Szakal, A., Nelson-Rees, W., Todaro, G. (1976) A clone isolated from HT-29 mucus-secreting cells. The 3

41. Graham, F. L., Smiley, J., Russell, W. C., and Nairn, R. (1977) Characterization of a new primary human pancreatic tumor line. Proc. Natl. Acad. Sci. U.S.A. 74, 335–345

42. Cotton, C. U., and Harris, A. (2009) Intronic enhancers coordinate epithelial-specific looping of the active CFTR locus. Proc. Natl. Acad. Sci. U.S.A. 106, 19934–19939

43. Wendt, K. S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., Yahata, K., Imamoto, F., Aburutani, H., Nakao, M., Imamoto, N., Maeshima, K., Shirahige, K., and Peters, J. M. (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature 451, 796–801

44. Hagege, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Fornè, T. (2007) Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat. Protoc. 2, 1722–1733

45. Blackledge, N. P., Ott, C. J., Gillen, A. E., and Harris, A. (2009) An insulator element 3' to the CFTR gene binds CTCF and reveals an active chromatin hub in primary cells. Nucleic Acids Res. 37, 1086–1094

46. Dohrmann, A., Miyata, S., Gallup, M., Li, J. D., Chapelin, C., Coste, A., Escudier, E., Nadel, J., and Basbaum, C. (1998) Mucin gene (MUC2 and MUC5AC) upregulation by Gram-positive and Gram-negative bacteria. Biochim. Biophys. Acta 1406, 251–259

47. Smirnova, M. G., Guo, L., Birchall, J. P., and Pearson, J. P. (2003) LPS up-regulates mucin and cytokine mRNA expression and stimulates mucin and cytokine secretion in goblet cells. Cell. Immunol. 221, 42–49

48. Hauber, H. P., Goldmann, T., Vollmer, E., Wollenberg, B., and Zabel, P. (2007) Effect of Dexamethasone and ACC on Bacteria-Induced Mucin Expression in Human Airway Mucosa. Am. J. Respir. Cell Mol. Biol. 37, 606–616

49. Li, T., Hu, J. F., Qiu, X., Ling, J., Chen, H., Wang, S., Hou, A., Yu, T. H., and Hoffman, A. R. (2008) CTCF regulates allelic expression of Ifg2 by orchestrating a promoter-polycomb repressive complex 2 intrachromosomal loop. Mol. Cell Biol. 28, 6473–6482

50. Lutz, M., Burke, L. J., Barreto, G., Goeman, F., Greb, H., Arnold, R., Schultheiss, H., Brehm, A., Kouzarides, T., Lobanenko, V., and Renkawitz, R. (2000) Transcriptional repression by the insulator protein CTCF involves histone deacetylases. Nucleic Acids Res. 28, 1707–1713

51. Yamada, N., Hamada, T., Goto, M., Tsutsumida, H., Higashi, M., Nomoto, M., and Yonezawa, S. (2006) MUC2 expression is regulated by histone H3 modification and DNA methylation in pancreatic cancer. Int. J. Cancer 119, 1850–1857

52. Hauber, H. P., Goldmann, T., Vollmer, E., Wollenberg, B., and Zabel, P. (2007) Effect of Dexamethasone and ACC on Bacteria-Induced Mucin Expression in Human Airway Mucosa. Am. J. Respir. Cell Mol. Biol. 37, 606–616

53. Li, T., Hu, J. F., Qiu, X., Ling, J., Chen, H., Wang, S., Hou, A., Yu, T. H., and Hoffman, A. R. (2008) CTCF regulates allelic expression of Ifg2 by orchestrating a promoter-polycomb repressive complex 2 intrachromosomal loop. Mol. Cell Biol. 28, 6473–6482

54. Lutz, M., Burke, L. J., Barreto, G., Goeman, F., Greb, H., Arnold, R., Schultheiss, H., Brehm, A., Kouzarides, T., Lobanenko, V., and Renkawitz, R. (2000) Transcriptional repression by the insulator protein CTCF involves histone deacetylases. Nucleic Acids Res. 28, 1707–1713

55. Yamada, N., Hamada, T., Goto, M., Tsutsumida, H., Higashi, M., Nomoto, M., and Yonezawa, S. (2006) MUC2 expression is regulated by histone H3 modification and DNA methylation in pancreatic cancer. Int. J. Cancer 119, 1850–1857

56. Li, D., Gallup, M., Fan, N., Szymkowski, D. E., and Basbaum, C. B. (1998) Cloning of the Amino-terminal and 5' Flanking Region of the Human MUC5AC Mucin Gene and Transcriptional Up-regulation by Bacterial Exoproducts. J. Biol. Chem. 273, 6812–6820

57. Li, J. D., Dohrmann, A. F., Gallup, M., Miyata, S., Gunn, J. R., Kim, Y. S., Nadel, J. A., Prince, A., and Basbaum, C. B. (1997) Transcriptional activation of mucin by Pseudomonas aeruginosa lipopolysaccharide in the pathogenesis of cystic fibrosis lung disease. Proc. Natl. Acad. Sci. U.S.A. 94, 967–972