Mucins and inflammatory bowel disease

Tarek Shirazi, Robert J Longman, Anthony P Corfield, Christopher S J Probert

Abstract
There is a layer of mucus lining the gastrointestinal tract, which acts as both a lubricant and as a physical barrier between luminal contents and the mucosal surface. The mucins that make up this layer consist of a protein backbone with oligosaccharides attached to specific areas of the protein core. These areas are called the variable number tandem repeat regions. The degree of glycosylation of the mucins is central to their role in the mucus barrier. The oligosaccharides are variable and complex. It has been demonstrated that the degree of sulphation and sialylation and the length of the oligosaccharide chains all vary in inflammatory bowel disease. These changes can alter the function of the mucins. Mucins are broadly divided into two groups, those that are secreted and those that are membrane bound. The major mucins present in the colorectum are MUC1, MUC2, MUC3, and MUC4.

Trefoils are a group of small peptides that have an important role in the mucus layer. Three trefoils have been demonstrated so far. They seem to play a part in mucosal protection and in mucosal repair. They may help to stabilise the mucus layer by cross linking with mucins to aid formation of stable gels. Trefoils can be expressed in the ulcer associated cell lineage, a glandular structure that can occur in the inflamed mucosa. There seem to be differences in the expression of trefoils in the colon and the small bowel, which may imply different method of mucosal repair.

A layer of mucus, which seems to serve two purposes, lines the gastrointestinal tract. It acts firstly as a lubricant and secondly as a protective physical barrier between the mucosal surface and the luminal contents. This layer is formed by mucus glycoproteins or mucins. These are huge molecules, typically with a molecular mass of the order of $10^6$ Daltons. By way of comparison, albumin has a molecular weight of $69 \times 10^3$ Daltons. These are huge molecules, typically with a molecular mass of the order of $10^6$ Daltons. By way of comparison, albumin has a molecular weight of $69 \times 10^3$ Daltons. They consist of a central protein backbone with large numbers of attached oligosaccharides. The protein core consists of two distinct types of amino acid sequence, heavily glycosylated regions and relatively sparsely glycosylated areas. The latter areas include cysteine rich domains similar to the D-domains in von Willebrand factor, thought to be important for dimer and oligomer formation. Common to all mucin protein cores that have so far been identified are tandem repeat sequences or variable number tandem repeats (VNTR). These are sequences of amino acids which are repeated and contain a high proportion of serine and threonine, the attachment sites for O-linked oligosaccharides. Accordingly, these are the regions that are heavily glycosylated (figs 1 and 2). Glycosylation accounts for up to 60%–80% of the mass of the molecule, and is responsible for many of the properties of mucins. It consists principally of oligosaccharides in O-linkage of N-acetylgalactosamine to serine or threonine in the protein core. This review aims to look at the evidence that a defect in mucins may contribute to mucosal diseases such as inflammatory bowel disease (IBD).

Mucin glycosylation
Central to the role of mucosal protection is the degree of glycosylation of the mucins. The VNTR segments of the central protein core...
MUC1 and cytoplasm. Figure 2

Cytoplasm

MUC1

Membrane anchor

Cytoplasmic tail

Approximate representation of MUC1 at an apical surface membrane

Lumen

10 mm

Glycocalyx

Membrane anchor

Cytoplasmic tail

Plasma membrane

have carbohydrate chains attached in a “bottle brush” fashion. Colonic mucins normally have up to 12 monosaccharides per chain. The oligosaccharides found in individual mucins are typically variable and can be complex. The carbohydrate structures themselves can be either linear or branched, and can be acidic (containing sialic acid or sulphate groups) or neutral in nature. There is, therefore, a huge scope for variation.

In general, the oligosaccharides are themselves made of three separate regions. There is a core sequence attached to the peptide, a backbone sequence (which may or may not be present) and finally, a peripheral region which is made up of sialic acid, blood group, or ester sulphate substitutions. The patterns of glycosylation are tissue specific within the gastrointestinal tract. The formation of oligosaccharide chains is governed by a series of glycosyltransferases each of which is specific for each link of the chain.

Work from Pullan and colleagues has shown that there is a variation in the thickness of the mucin layer between disease groups. The layer is thinner than normal in ulcerative colitis and thicker than normal in Crohn’s disease. The reduction of the thickness in ulcerative colitis is felt to reflect the goblet cell depletion found in ulcerative colitis. Goblet cell depletion is striking by its absence in Crohn’s disease. It is not at present known why there is retention of goblet cells in Crohn’s disease since in all other cases where there is acute inflammation of the colon, there is goblet cell depletion.

Mucin biochemistry

There have been a number of important biochemical changes noted in the mucins in IBD such as glycosylation and sulphation. The oligosaccharide chain length in IBD has been shown to be half that of the normal state. Sulphation is reduced and sialylation is increased. There is also increased expression of short oncofetal carbohydrate structures, such as galactose 1,3 N-acetylgalactosamine. The changes described are likely to alter the viscoelastic properties of the gels formed and influence interactions of mucins with microorganisms, electrolytes, defensive proteins and dietary components, hence reducing the effectiveness of supramucosal layer function. Sulphation and sialylation are important as they play a part in the resistance of mucins to bacterial degradation. Sialylation is controlled by a family of sialyl transferases, specific for the different linkages of sialic acid found in mucin oligosaccharides. In addition, colonic mucin sialic acids are O-acetylated, in keeping with their role in resistance to degradation. It has long been assumed that these changes are secondary to the disease itself, being a result of either quicker passage through the Golgi, or altered post-translational processes. One study has suggested a possible racial variation. Asians with colitis do not exhibit the reduction in sulphation seen in their European counterparts (figs 3 and 4). Moreover, Asian colitics do not have the same increase in incidence of colonic carcinoma. Indeed, they seem to have a benign form of the disease since not only does their mortality remain low, but also they have significantly lower resection rates. This suggests genetically determined glycosylation abnormalities.

The factors outlined all suggest that the effectiveness of the mucus barrier is reduced in inflammatory bowel disease. This may make it more susceptible to bacterial degradation. Enzymatic desulphation by faecal bacterial sulphatases leads to an increased susceptibility to degradation by faecal glycosidases. Work has shown that about 1% of normal colonic bacteria secrete a range of enzymes which completely degrade mucin oligosaccharides, the remaining numerically dominant strains possess only some of these enzymes. This allows the total enteric microflora to use mucin carbohydrate as an energy source. The mucus layer is in equilibrium between synthesis and bacterial degradation by the colonic microflora. The degree of sulphation in colonic mucins is far greater than that seen in mucins elsewhere (for example stomach or small intestine). Sulphation of mucins is also seen to correlate with the presence of bacteria and this may represent an adaptive response. It has been shown that there is increased mucin sulphatase activity in ulcerative colitis, particularly in active rather than quiescent disease. It was also demonstrated that the faecal sulphatase activity mirrors the disease activity. Whether this is cause or effect is
unclear at this time. It is easy to see how the balance can be tipped away from a stable gel barrier to a situation where the mucus layer is degraded.

**Mucin genes**

Nine genes (MUC genes) have thus far been identified which code for the protein cores of mucins. There are, broadly speaking, two groups of mucins. The first are secreted mucins, which are the gel forming mucins found on the mucosal surface. Secondly, there are the membrane associated mucins found at the apical membranes of the epithelial cells. The major genes coding for the gel forming mucins are situated, as a cluster, on chromosome 11.p15.5. These genes are MUC2, MUC5AC, MUC5B, and MUC6. Until recently, it was assumed that only MUC1 existed in a membrane bound form. It is now known that both MUC3 and MUC4 may also exist in membrane bound forms. The predominant mucins detected in the colorectum are MUC1, MUC2, MUC3, and MUC4. MUC2 is normally specific to goblet cells. Further evidence for separate organisation of secreted and membrane associated MUC genes has come from developmental studies. MUC3 and MUC4 appear at 6.5 weeks of gestation, a stage when the epithelium is stratified and undifferentiated, as part of the primitive gut. MUC1 appears in the colon at 18 weeks of gestation. The major secreted mucin, MUC2, is first expressed later than any of these three genes (that is MUC 1, 3, and 4). These are found in mucus secreting and non-mucus secreting epithelial cells and are

---

**Figure 3** (A) Shows normal colon and (B) the loss of sulphation seen in European colitics; contrast this with fig 4.

**Figure 4** (A) Shows Asian control and (B) the retention of sulphation seen in Asian colitics; contrast this with fig 3.
clearly implicated in cellular roles reflecting their association with membranes. The adult pattern of major MUC2 and MUC4 gene expression and low to background MUC1 and MUC3 predominates from birth onwards. Investigation into the expression of the MUC genes in IBD has been carried out at the nucleic acid or transcription level using quantitative and semiquantitative techniques of mRNA analysis. Additionally, antibodies to the VNTR and the non-VNTR regions of the MUC peptides have been used to detect the translated mucin products. In ulcerative colitis, mRNA levels seem to be very similar to normal controls with strong expression of MUC2 and MUC4. There are lower levels of MUC1 and MUC3. This pattern is not significantly different from normal. A recent study has shown an association between rare alleles of the MUC3 gene and ulcerative colitis. In Crohn’s disease, normal levels of MUC2 and MUC3 have been reported. In a more detailed study of MUC gene expression, MUC2, MUC3, and MUC4 were all strongly expressed in the ileum, in normal and Crohn’s disease tissue. This study showed a reduction in MUC1 in inflamed ileum when compared with areas of normal mucosa from the same patient. When compared with normal controls, there was also a reduction in MUC3, MUC4, and MUC5B (detected as very low amounts) levels in patients with Crohn’s disease. These particular changes were found in both the diseased and the normal tissues from the patients with Crohn’s disease. This implies an early or primary mucosal defect in Crohn’s disease, as opposed to the MUC1 changes that seem to be related to inflammation. In a model of the adenoma-carcinoma sequence, there is upregulation of MUC2 and neoexpression of MUC5AC in adenomas. In contrast, de novo carcinoma appears to be associated with loss of mucin secretion. The increased expression of MUC2 appears to correlate with a poorer survival rate.

Trefoils
A group of small cysteine rich peptides, trefoil peptides, may also have an important role in the mucus layer. Three human trefoil peptides have been discovered to date. In an effort to standardise nomenclature, these are now referred to as trefoil factor family (TFF) peptides. Normally, trefoil peptides seem to be expressed in a site specific fashion. TFF1 is found in the foveolar cells of the stomach, and TFF2 in the distal stomach and lower portion of Brunner’s glands of the duodenum; TFF3 is found throughout the small and large intestine.

Trefoils seem to have two important roles, that of epithelial protection and that of mucosal healing. They may play a part in mucus stabilisation, by interacting or cross linking with mucins to aid formation of the gel layer. When mucosal injury occurs trefoils are rapidly upregulated and stimulate repair by a process known as epithelial restitution. Trefoil peptides may be coexpressed with secreted mucins. Unpublished data suggest a possible synergistic action in mucosal protection and repair between the chromosome 11 mucins and trefoil peptides, which are coexpressed in both normal and diseased mucosa. TFF2 in the stomach seems to alter acid permeation through the gel layer. This could act by reducing acid backflow through the gel from lumen to epithelium, thus protecting the cells from acid damage. Another study showed that overexpression of TFF1 in transgenic mice increases their resistance to epithelial injury induced by non-steroidal anti-inflammatory drugs (NSAIDs). In rats it has been demonstrated that there is less damage to the epithelium of the stomach due to NSAIDs, alcohol or restraint stress when exogenous TFF2 is administered.

The part played by trefoils in mucosal healing is the subject of ongoing research. In a number of animal models and cell lines, trefoils promote mucosal healing and cell migration across an area of injury (restitution) respectively. One study looking at TFF3 “knock-out mice” demonstrated an impairment of colonic epithelial healing after oral dextran sulphate challenge, when compared with wild-type mice who develop a mild and transient colitis. Repletion of TFF3 leads to an improvement of the severe induced colitis in the knock-out mice.
Ulceral associated cell lineage

In many instances of gastrointestinal inflammation, a potentially reparative glandular structure can occur in the mucosal layer. Wright et al first described this as the ulceral associated cell lineage (UACL) in 1990, in relation to the ulcers found in small bowel Crohn's disease. UACL-like phenomena have subsequently been described in Barrett's oesophagus, duodenal ulcers, atrophic gastritis, as well as small bowel Crohn's disease. UACL is a stem cell lineage that has the ability to express all three trefoils as well as epidermal growth factor and lysoyzme. There is a zone of proliferation, which provides a constant supply of cells that can migrate onto the luminal surface of the mucosa. This acts to cover the epithelial breach. UACL may therefore play a central part in epithelial repair and this is driven by trefoil peptides. There is, however, no evidence of ectopic TFF1 or TFF2 in large bowel inflammation and there is reduced TFF3 expression in ulcerative colitis. This suggests that there are different methods of mucosal repair between the colon and the small bowel.

There is evidence that trefoil peptides, in particular TFF2, promote migration of cells through collagen gels (although this work used breast cancer cell lines rather than gastrointestinal inflammation). This is important as, after injury, there is a rapid covering of the denuded area by fibrin and necrotic tissue through which cells have to migrate. There is some evidence that the prior administration of trefoils can lessen gastric injury induced by indomethacin. With all this evidence, a therapeutic role based on trefoil peptides is possible in cases of mucosal repair between the colon and the small bowel.

The future

There are changes in the structure of the colonic mucins, which may affect their function as a protective barrier. These differences relate mainly to the carbohydrate side chains that are bound to the protein core. The changes are different in ulcerative colitis and Crohn's disease, for example, the thickness of the gel layer and the degree of sulphation. Gene expression in ulcerative colitis is not significantly different from normals, while changes in ileal Crohn's disease mucin gene expression are found. Work is needed to look further at the nature of the mucin products synthesised and secreted in Crohn's colitis, and ulcerative colitis. Improved antinucin antibodies are now becoming available to address this question. In addition, the work looking at Asian colitics and the sulphation changes in mucin alluded to earlier, can now be backed up by specifically looking at any variations in MUC gene expression.

The most significant improvements in mucin analysis that have taken place in the last 2-3 years have been the ability to separate and detect very small quantities of native mucins with precise MUC gene identity. Following on from the wave of molecular biological data giving greater definition of the family of MUC genes, our understanding of the mucins in general is now at a level sufficient to address the fundamental roles played in the inflammatory and other mucosal diseases. It remains to be seen whether the evidence of mucin gene discovery. It is also clear that the scope for interaction in many defensive, reparatory (for example the trefoil peptides) and other mucosal systems exists through the abundance of individual glycosylation patterns associated with mucins throughout the gastrointestinal tract. Although this poses a continual analytical problem due to the huge variety of oligosaccharide structures, it is clear that much will be gained from the further study of mucin glycobiology. Understanding mucins is now a valid aim in studies on IBD and promises real advances in the immediate future.

The work reported in this review was supported by grants from the Wellcome Trust (grant 051886/Z/97), Smith and Nephew, the Royal College of Surgeons, and the South and West Regional Health Authority.

References

Shirazi, Longman, Corfield, et al


