



HAL
open science

The mouse *Muc5b* mucin gene is transcriptionally regulated by thyroid transcription factor-1 (TTF-1) and GATA-6 transcription factors.

Nicolas Jonckheere, Amélie Velghe, Marie-Paule Ducourouble, Marie-Christine Copin, Ingrid B. Renes, Isabelle van Seuningen

► To cite this version:

Nicolas Jonckheere, Amélie Velghe, Marie-Paule Ducourouble, Marie-Christine Copin, Ingrid B. Renes, et al.. The mouse *Muc5b* mucin gene is transcriptionally regulated by thyroid transcription factor-1 (TTF-1) and GATA-6 transcription factors.: Regulation of *Muc5b* mucin gene by TTF-1 and GATA factors. *FEBS Journal*, 2011, 278 (2), pp.282-94. 10.1111/j.1742-4658.2010.07945.x . inserm-00807850

HAL Id: inserm-00807850

<https://inserm.hal.science/inserm-00807850>

Submitted on 4 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



The mouse Muc5b mucin gene is transcriptionally regulated by TTF-1 and GATA-6 transcription factors

Journal:	<i>FEBS Journal</i>
Manuscript ID:	FJ-10-0798.R1
Manuscript Type:	Regular Paper
Subdiscipline:	Gene expression, transcription and translation
Date Submitted by the Author:	20-Oct-2010
Complete List of Authors:	Jonckheere, Nicolas; INSERM U837/JPARC, Team 5 " mucins, epithelial differentiation and carcinogenesis" Velghe, Amélie; INSERM U837/JPARC, Team 5 "mucins, epithelial differentiation and carcinogenesis" Ducourouble, Marie-Paule; INSERM U837/JPARC, Team 5 "mucins, epithelial differentiation and carcinogenesis" Copin, Marie-Christine; INSERM U837/JPARC, Team 5 " mucins, epithelial differentiation and carcinogenesis" Renes, Ingrid; Erasmus MC and Sophia Children's Hospital, Division of Neonatology, Department of Pediatrics van Seuning, Isabelle; INSERM U837/JPARC, Team 5 " mucins, epithelial differentiation and carcinogenesis"
Key Words:	mucin, Muc5b, TTF1, GATA, lung, differentiation

1
2
3 **The mouse *Muc5b* mucin gene is transcriptionally regulated by TTF-1 and GATA-6**
4
5 **transcription factors**
6
7
8
9

10 Nicolas Jonckheere^{1,2}, Amélie Velghe¹, Marie-Paule Ducourouble^{1,2}, Marie-Christine Copin¹,
11
12 ^{2,3}, Ingrid B. Renes⁴, and Isabelle Van Seuning^{1,2}
13
14
15
16

17 ¹Inserm, U837, Team #5 “Mucins, epithelial differentiation and carcinogenesis”, rue
18 Polonovski, 59045 Lille Cedex, France
19
20

21
22 ²Université Lille Nord de France, 1 Place de Verdun, 59045 Lille cedex, France
23

24
25 ³Centre de Biologie-Pathologie, Centre Hospitalier Régional et Universitaire de Lille,
26 Boulevard du Professeur Jules Leclercq, 59000 Lille, France
27
28

29 ⁴Laboratory of Pediatrics, Division of Neonatology, Erasmus MC-Sophia Hospital,
30 Rotterdam, The Netherlands
31
32
33
34
35

36 Running title: Regulation of *Muc5b* mucin gene by TTF-1 and GATA factors
37
38
39
40

41 Genbank accession number : AY744445
42
43
44
45

46 To whom correspondence should be sent:
47

48 Nicolas Jonckheere, Ph.D.
49

50 ¹Inserm, U837, Team #5, Rue Polonovski, 59045 Lille cedex, France
51
52

53 Phone: 33.320.29.88.67
54

55 FAX: 33.320.53.85.62
56

57 E-mail: nicolas.jonckheere@inserm.fr
58
59
60

Abstract

MUC5B is one of the major mucin genes expressed in the respiratory tract. Previous studies in our laboratory have demonstrated that *MUC5B* is expressed in human lung adenocarcinomas and during lung morphogenesis. Moreover, in human lung adenocarcinoma tissues, a converse correlation between *MUC5B* and TTF-1 expression, a lung-specific transcription factor, was established. However, the molecular mechanisms that govern the regulation of *MUC5B* expression in the lung are largely unknown. In order to better understand the biological role of *MUC5B* in lung pathophysiology, we report now the characterization of the promoter region of the mouse *Muc5b* mucin gene. The promoter is flanked by a TATA box (TACATAA) identical to that in the human gene. Human and murine promoters share 67.5% similarity over the first 170 nucleotides. By RT-PCR, co-transfection studies and gel-shift assays we show that *Muc5b* promoter activity is completely inhibited by TTF-1 whereas factors of the GATA family (GATA-4/-5/-6) are activators. Altogether, these results demonstrate for the first time that *Muc5b* is a target gene of transcription factors (TTF-1, GATA-6) involved in lung differentiation programs during development and carcinogenesis and identifies TTF-1 as a strong repressor of *Muc5b*. The characterization of the structural and functional features of *Muc5b* mucin gene will provide us with a strong base to develop studies in murine models aimed at identifying its biological role in lung pathophysiology.

Keywords: mucin, *Muc5b*, TTF1, GATA, lung, differentiation

Introduction

Mucins are high molecular weight glycoproteins that are synthesized by specialized epithelial cells and are thought to promote tumor cell invasion [1]. In the tracheobronchial tree, the main mucin genes are *MUC5B* and *MUC5AC*, that encode two secreted mucins, and *MUC4*, that encodes a transmembrane mucin [2]. *MUC5B* and *MUC5AC* are expressed in mucus-producing cells with *MUC5AC* in the surface goblet cells and *MUC5B* in the mucous cells of the submucosal gland whereas *MUC4* is found in a wide array of epithelial cells [3-7].

MUC5B expression, in the developing lung, is seen as of 13 weeks of gestation in the epithelial folds of the surface epithelium [8]. At a later stage, *MUC5B* is found in cells of the gland ducts and mucous glands [9]. In the adult lung, the expression of *MUC5B* follows a restricted pattern with a positive gradient from the surface to the glands and a decrease in intensity from the tracheo-bronchus toward the bronchioles, with no signal in small bronchioles and pneumocytes [10]. The murine *Muc5b* mucin gene was recently characterized in our laboratory and was shown to be expressed in mucous cells of the laryngeal glands [11]. In lung adenocarcinomas, *MUC5B* is frequently expressed in mucus-secreting carcinomatous cells [12]. In mucinous type of bronchioloalveolar carcinoma, *MUC5B* expression is the most intense along with that of *MUC5AC*. The expression of *MUC5B* is lost in poorly differentiated and non mucinous lung carcinomas [12]. From these studies, it appears that *MUC5B* may be used as a marker of cytodifferentiation in the lung associated with mucous differentiation [5].

The early expression of mucin genes before mucous cell differentiation or during the process of differentiation suggests that they may be the targets of transcription factors responsible for those programs [13]. Accordingly to that hypothesis, we recently showed that *MUC2* and *MUC4* are transcriptionally regulated by Cdx homeodomain proteins and GATA

1
2
3 factors [14-16]. The thyroid transcription factor-1 (TTF-1) is an important factor during lung
4
5 morphogenesis [17-20] and drives the expression of several lung-specific genes such as
6
7 surfactant proteins [21], CCSP [22] and CC10 [23]. Moreover, we have recently shown a
8
9 converse correlation between MUC5B and TTF-1 expression in human lung
10
11 adenocarcinomatous tissues [24], suggesting a negative regulation of *MUC5B* by this
12
13 transcription factor. GATA factors also possess a restricted pattern of expression during lung
14
15 development [25]. GATA-6 seems to be involved during different phases of the development
16
17 [26] whereas GATA-5 plays a role in transcriptional programs in the earliest steps of lung
18
19 development [27]. Moreover, synergistic mechanisms between homeoprotein TTF-1 and zinc-
20
21 finger GATA-6 have been recently described [21,28].
22
23
24
25
26

27 Having found (i) binding sites for these factors in both the human [29] and murine (this
28
29 report) *MUC5B* mucin genes, (ii) a restricted pattern of *MUC5B* expression in the respiratory
30
31 tract [5] and (iii) expression of MUC5B, TTF-1 and GATA-6 in lung adenocarcinomas
32
33 [12,24,30,31], we undertook to study the regulation of the *Muc5b* promoter by TTF-1 and
34
35 GATA factors. By this approach we aimed to show transcriptional regulation of *Muc5b* by
36
37 these two transcription factors thereby providing a strong base for the development of studies
38
39 aimed at identifying the biological role of *Muc5b* in the lung using mouse models.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Results

Characterization of the sequence of the promoter of the murine Muc5b mucin gene

The sequence covering 1210 nucleotides upstream of the transcription initiation site is shown in Fig. 1A. It is characterized by the presence of a TATA box (TACATAA) at -28/-22. The immediate sequence is GC-rich and contains a few putative binding sites for Sp1-like factors (GC-boxes and CACCC boxes). We also note the presence of putative binding sites for the lung-specific factor TTF-1 throughout the sequence. GATA putative binding sites are present both in the proximal and distal parts of the promoter.

Alignment of the human and mouse promoter sequences showed that there is a high homology (67.5 %) over the first 157 nucleotides flanking the TATA box (Fig. 1B) and that the sequence of the TATA box (TACATAA) is identical in the two species.

Characterization of Muc5b promoter activity

Mouse *Muc5b* transcriptional regulation at the promoter and mRNA levels was studied in the murine CMT-93 colorectal cancer cell line that is commonly used to study murine mucin gene regulation since it is known to express several mucin genes [15,32] and of interest in this study expresses *Muc5b* mRNA (Fig. 2A). Since no murine lung epithelial cell line expressing *Muc5b* is available at this time we also studied *mMuc5b* promoter regulation in the human lung NCI-H292 cell line that expresses *MUC5B* as we previously demonstrated [33]. To define essential regions that drive transcription of the *Muc5b* promoter, six deletion mutants, that cover 1.2 kb of the promoter, were constructed in the promoterless pGL3 basic vector (Fig. 2B). Data indicate that the promoter is active both in murine intestinal CMT-93 and human lung NCI-H292 cell lines. The four deletion constructs tested (-169/-1, -478/-1, -717/-1 and -1195/-1) have similar luciferase activities in each cell line, which suggests that

1
2
3 the proximal region -169/-1 is sufficient to drive maximal activity of the promoter in these
4
5 cells (Fig. 2C). The influence of the 5'-UTR on the promoter activity was studied using the
6
7 constructs -478/+47 and -717/+47. When the 5'-UTR region +1/+47 was included, the
8
9 activity of the promoter remained similar (compare activities of -478/-1 to -478/+47 and of -
10
11 717/+47 to -747/-1).
12
13
14
15
16

17 ***TTF-1 is a strong repressor of Muc5b expression***

18
19 Overexpression of TTF-1 in CMT-93 cells led to a strong decrease of the amount of
20
21 *Muc5b* mRNA (75 % loss, Fig. 3A). Co-transfection experiments in the presence of the
22
23 pCMV-TTF-1 expression vector showed that overexpression of TTF-1 also led to a dramatic
24
25 decrease (60-75 %) of the activity of *Muc5b* promoter both in CMT-93 and NCI-H292 cells
26
27 (Fig. 3B). The decrease was even more pronounced in NCI-H292 cells (80 % loss). The
28
29 strong inhibition was seen with all the constructs tested in this work suggesting that the -477/-
30
31 1 region is sufficient to convey the repression of *Muc5b* promoter by TTF-1. TTF-1 binds to
32
33 the -CAAG- consensus sequence. Putative binding sites were found throughout the sequence
34
35 of the *Muc5b* promoter (see Fig. 1A). EMSAs were performed with several probes containing
36
37 TTF-1 consensus binding sites found in the murine promoter (Table 2) as well as with their
38
39 mutated version (CAAG to **GTAT**). The probe T211 contains two putative TTF-1 binding
40
41 sites at -358/-355 and -353/-350 and the probe T212 contains two sites at -709/-706 and -
42
43 700/-697, respectively. The probes T213, T238 and T242 contained one predicted site at -
44
45 112/-109, -325/-322, and -417/-414, respectively. Incubation of T211 and T212 radiolabeled
46
47 probes with nuclear proteins from NCI-H292 cells produced one specific shifted band (Fig.
48
49 3C, lanes 2 and 8). Specificity of the complex was confirmed by the loss of the shifted band
50
51 (indicated by an asterisk) when cold probes, in a 50x excess, were incubated with nuclear
52
53 proteins before adding the radiolabeled probe (lanes 3 and 9). Moreover, no competition
54
55
56
57
58
59
60

1
2
3 could be observed when mutated probes were used in the competition (lanes 4 and 10).
4
5 Implication of TTF-1 in the complex formation was further confirmed when mutated probes
6
7 were radiolabeled and incubated with nuclear extracts. In that case, no binding was visualized
8
9 (lanes 6 and 12). The probe T238 did not produce any shift and the probes T213 and T242
10
11 that contained a predicted TTF-1 site did not bind TTF-1 (not shown).
12
13
14
15

16 17 ***Role of GATA factors in the regulation of Muc5b expression*** 18

19
20 Besides TTF-1, GATA factors and especially GATA-6 are important factors in lung
21
22 morphogenesis and are known to (i) regulate TTF-1 and (ii) synergize with TTF-1 to activate
23
24 transcription of their target genes. Analysis of the sequence of the promoter of *Muc5b* showed
25
26 that putative binding sites for GATA factors were present throughout the sequence (see Fig.
27
28 1A), which is in favor of a possible role in the regulation of *Muc5b*.
29
30

31
32 At the mRNA level, we observed an increase of *Muc5b* expression with GATA-5 (4
33
34 fold) and GATA-6 (14 fold) once we had overexpressed those transcription factors in CMT-
35
36 93 cells (Fig. 4A). There was no effect visualized with GATA-4. To localize the GATA
37
38 responsive elements we then performed co-transfection experiments both in the CMT-93 (Fig.
39
40 4B) and NCI-H292 (Fig. 4C) cell lines. Overexpression of GATA-5 in CMT-93 cells induced
41
42 a strong activation of the three constructs of the *Muc5b* promoter (4-, 4- and 6-fold activation
43
44 on -478/-1, -717/-1 and -1195/-1 constructs, respectively, $p < 0.05$). Overexpression of GATA-
45
46 4 and GATA-6 in these cells also induced transactivation of -717/-1 and -478/-1 *Muc5b*
47
48 promoter constructs, respectively (2-4 fold activation, $p < 0.05$) (Fig. 4B). In lung NCI-H292
49
50 cells, the profile was slightly different in that the strong transactivating effect of the three
51
52 GATA factors on the -717/-1 region went down with the -1195/-1 deletion construct (Fig.
53
54 4C). This suggests that some inhibitory factors binding to the -1195/-718 region of the
55
56 promoter may interfere with GATA function in these cells.
57
58
59
60

1
2
3 GATA *cis*-elements within the promoter of *Muc5b* were then identified by performing
4 EMSA experiments with DNA probes containing GATA putative binding sites located at -
5 411/-408 (T242), -454/-449 (T254) and -1143/-1140 (T84), respectively. As shown in Fig.
6
7
8 4D, incubation of these three probes with nuclear proteins from CMT-93 cells produced one
9
10 specific shifted complex (GATA) (lanes 2, 9 and 15, respectively). Specificity was confirmed
11
12 by complete inhibition of the complex formation when unlabelled competition was performed
13
14 with a 50x excess of the cold probe (lanes 3, 10 and 16). GATA-4 and GATA-6 were both
15
16 able to bind the GATA element present in T254 and T84 probes since a supershift was
17
18 visualized upon addition of a GATA-4 (lanes 11 and 17) or GATA-6 (lanes 13 and 19)
19
20 antibody in the mixture. GATA-4 is involved in the complex formation with the T242 probe
21
22 since a supershift was observed upon addition of the anti-GATA-4 antibody in the reaction
23
24 mixture (lane 5). No supershift was seen when anti-GATA-5 antibody was used (lanes 6, 12
25
26 and 18). However, we can not conclude that this factor does not bind to these sites since it
27
28 also did not induce a supershift when a commercial consensus GATA probe was used (not
29
30 shown). ChIP assay was carried out on -503/-261 region of *Muc5b* promoter containing
31
32 notably T242 (-411/-408) and T254 (-454/-449) binding sites. Binding of GATA-4, GATA-5
33
34 and GATA-6 to the *Muc5b* promoter was observed in CMT-93 cells (Fig. 4E). Specificity of
35
36 the binding was confirmed by complete absence of PCR amplification using IgGs.
37
38
39
40
41
42
43
44

45
46 In order to show a possible synergistic mechanism of regulation between TTF-1 and
47
48 GATA-6, co-transfections with those two factors were carried out in CMT-93 cells with the
49
50 -478/-1, -717/-1 and -1195/-1 *Muc5b* promoter constructs (Fig. 4E). As previously shown,
51
52 overexpression of GATA-6 transactivated the three deletion mutants, whereas overexpression
53
54 of TTF-1 strongly repressed the transcriptional activity of the three constructs. When co-
55
56 transfected together, TTF-1 inhibited the transactivating effect of GATA-6, which led to a
57
58
59
60

1
2
3 loss of the transactivation of the *Muc5b* promoter. The same result was obtained in NCI-H292
4
5 cells (not shown).
6
7
8
9

10 ***Expression of MUC5B, TTF-1 and GATA-6 in well-differentiated mucus-secreting lung***
11 ***adenocarcinomas***
12
13

14
15 Immunohistochemical analyses revealed that in well-differentiated mucus-secreting
16 lung adenocarcinomas, MUC5B expression was intense and cytoplasmic (Fig. 5A), whereas
17 there was no expression of TTF-1 in MUC5B-positive cells (Fig. 5B). In a papillary
18 adenocarcinoma, MUC5B was not detected (Fig. 5D). On the contrary, TTF-1 was expressed
19 in the nucleus of all these papillary adenocarcinomatous cells (Fig. 5E). In a non-mucinous
20 type of bronchioloalveolar carcinoma, TTF-1 was expressed in the majority of the
21 carcinomatous cells (Fig. 5H). In contrast, these TTF-1 positive cells did not express MUC5B
22 (Fig. 5H, inset). Interestingly, in another region of the same bronchioloalveolar carcinoma,
23 that was focally mucus-secreting, we found expression of MUC5B in a few mucus-secreting
24 tumor cells (Fig. 5G). In these MUC5B-expressing cells, TTF-1 was not expressed (Fig. 5G,
25 inset). Immunohistochemical studies on the same lung tumor tissues indicated that GATA-4
26 was not expressed in these samples (not shown) whereas GATA-6 was consistently expressed
27 in the cytoplasm of MUC5B-expressing cells (Fig. 5C, F and I).
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Discussion

Human *MUC5B* mucin gene is one of the main mucin genes expressed in the respiratory tract, in which it is mainly found in the mucous cells of the submucosal glands. Recently, we have characterized the human *MUC5B* promoter [29,34] and studied its expression both during lung development [8] and lung carcinogenesis [12]. From these studies it appears that the *MUC5B* promoter contains several putative binding sites for transcription factors playing critical roles in the formation, differentiation, and function of cells lining the respiratory tract such as TTF-1 and GATA factors [13]. Moreover, expression studies revealed a somewhat surprising early expression of *MUC5B* in the developing lung concomitant to mucous cell differentiation [8] and altered patterns of expression in lung adenocarcinomas [5,12,24].

The regulation of *MUC5B* by these transcription factors is however unknown and development of murine models of lung diseases are necessary to gain insight into, and understand, the regulation of the murine homologue of *MUC5B*. In the present study, we have isolated and characterized the promoter of the murine *Muc5b* mucin gene in order to study its transcriptional regulation by TTF-1 and GATA transcription factors. This approach will provide necessary knowledge to study *Muc5b* regulation in murine models and more particularly its biological role in lung pathophysiology.

Analysis of the promoter sequences of the murine *Muc5b* and the human *MUC5B* genes showed that they are highly similar over the first 170 nucleotides and more importantly that the TATA box is identical. This suggests that conserved regulatory mechanisms exist for these two genes throughout evolution and especially between mouse and human species.

Furthermore, in this report we demonstrate that TTF-1, which plays an important role in lung morphogenesis, lung repair after injury, or during carcinogenesis [20,35,36] and is a strong repressor of *Muc5b* expression at the promoter level. These results corroborate with our data in human tissues from different subsets of lung carcinoma in which we have also

1
2
3 shown a converse expression of TTF-1 and MUC5B proteins ([24] and this report) and with
4
5 another study that showed that mucinous parts of lung carcinomas expressing MUC5B are
6
7 TTF-1 negative [24,37]. Altogether, these results identify, for the first time, *Muc5b* as a direct
8
9 target gene of TTF-1, which most likely is responsible for the repression of *MUC5B* in certain
10
11 types of lung adenocarcinomas.
12
13

14
15 The main consequence of MUC5B repression by TTF-1 will be a modification in the
16
17 composition of the respiratory mucus since most of the mucus secretion in the lung comes
18
19 from mucous cells of the submucosal glands that secrete MUC5B [5,38]. The rheological
20
21 properties of the mucus and its ability to maintain a normal defense line against bacterial
22
23 infection, immune recognition of the cancer cell [1] or during development or repair [5] will
24
25 then be greatly impaired. In future studies, it will be interesting to determine whether
26
27 repression of MUC5B by TTF-1 represents a more general mechanism in lung diseases.
28
29
30

31
32 The GATA family of transcription factors is composed of several factors [25]. In the
33
34 lung it has been shown that GATA-4/-5/-6 are expressed in a restricted manner. These factors
35
36 participate in epithelial cell differentiation during embryonic development and establishment
37
38 of cell lineages derived from primitive intestine [39]. Previous work in our laboratory has
39
40 allowed identification of GATA factors as activators of mucin gene expression [15,16], such
41
42 as GATA-4 for Muc2 in intestinal cells [15], with obvious association between mucin
43
44 activation by GATA factors and terminal differentiation of the specialized epithelial cell in
45
46 which the mucin expression is activated. In this work, it appears that GATA-5 and GATA-6
47
48 are also activators of *Muc5b* transcription. GATA-4 only had a moderate effect on the
49
50 promoter activity. Previously, when we looked at GATA-4 expression in human lung tissues,
51
52 we could not find any expression of GATA-4. This is in agreement with a recent report that
53
54 showed that GATA-4 expression in lung carcinomas was repressed by hypermethylation of its
55
56 promoter [40]. Thus GATA-4 does not appear to be a candidate for *MUC5B* regulation in the
57
58
59
60

1
2
3 lung. Moreover, we consistently found positive cytoplasmic staining of GATA-6 in MUC5B-
4
5
6 expressing cells in the same sections used for TTF-1. A positive correlation was found
7
8 between GATA-6 and MUC5B expression in the same cells. However, despite the fact that
9
10 GATA-6 is a strong inducer of *Muc5b* transcription, its localization in the cytoplasm of
11
12 MUC5B-expressing lung carcinoma cells underscores its role as a major regulator of MUC5B
13
14 expression in the types of lung carcinomas studied in this report. Alteration of GATA-6
15
16 expression and aberrant cytoplasmic localization in ovarian cancer cells was recently
17
18 proposed to contribute to dedifferentiation of the tumor cells seen in the process of adaptation
19
20 to neoplastic progression [41].
21
22
23

24
25 Regulation of *Muc5b* expression by transcription factors expressed early during lung
26
27 development, such as TTF-1 and GATA-6, may have a critical role both in normal and
28
29 cancerous differentiation processes. From our data and others, the regulation of mucin genes
30
31 by GATA factors seems to be more general and may affect the expression of other mucin
32
33 genes, like *MUC2*, *MUC3* and *MUC4*, since their promoters also contain *cis*-elements for
34
35 these transcription factors [6,13,42,43]. Since GATA factors are expressed in endodermal
36
37 tissues, we hypothesize that this mechanism of regulation will occur in tissues derived from
38
39 endoderm and primitive gut including the lung but also the digestive tract as we have already
40
41 shown for *Muc2* expression by GATA-4 [15], *MUC4* by GATA-4/-5/-6 [16] and *MUC6* by
42
43 GATA-5/-6 (I Van Seuning, unpublished observations) in intestinal goblet cells. *MUC4*,
44
45 encodes a membrane-bound mucin expressed as early as 6.5 weeks after gestation, by the
46
47 primitive epithelial cells, and have the potential to differentiate in all the epithelial cell types
48
49 of the conducting airways and alveolar epithelium. In the lung, we think that *MUC4* may be a
50
51 good candidate [8]. TTF-1 and GATA-6 are required for the formation and differentiation of
52
53 distal epithelium [20,44,45]. In normal adult tissue, *MUC4* is preferentially expressed by the
54
55
56
57
58
59
60

1
2
3 epithelium of the tracheobronchial tract and is probably downregulated in alveolar cells [10].
4

5
6 Future studies have to confirm this hypothesis.
7
8
9

10 In conclusion, we have characterized the 5'-flanking region of the murine *Muc5b*
11 mucin gene and showed that the proximal part is highly homologous to its human counterpart.
12 We also showed that *Muc5b* is a direct target of and is transcriptionally regulated by the TTF-
13 1 (inhibitor) and GATA-6 (activator) transcription factors which are known to regulate cell
14 fate during lung morphogenesis. Altogether, the characterization of these structural and
15 functional features of the *Muc5b* mucin gene will allow studies in murine models
16 (inflammatory or cancerous) to define the biological role of Muc5b in lung pathophysiology.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Material and methods

Construction of *Muc5b-pGL3 deletion mutants*

The murine *Muc5b-pGL3* deletion mutants covering 1194 nucleotides upstream of the first ATG were constructed into pGL3 basic vector (Promega, Charbonnières-les-Bains, France) using a PCR-based method as previously described [29]. PCR reactions were carried out on Ali2 cosmid clone, previously used to isolate *Muc5b* 5'-flanking region [11]. Internal deletion mutants were generated by PCR using pairs of primers bearing specific restriction sites at their 5' and 3' ends (Table 1). PCR products were digested, gel-purified (QIAquick gel extraction kit, Qiagen, Courtaboeuf, France) and subcloned into the pGL3 Basic vector that had been previously cut with the same restriction enzymes. All clones were sequenced on both strands on an automatic LI-COR sequencer (ScienceTech, France) using infra-red labeled RV3 and GL2 primers (Promega). The promoter sequence was submitted to Genbank (accession number AY744445). Plasmids used for transfection studies were prepared using the Endofree plasmid Mega kit (Qiagen).

Cell culture

Murine rectal cancer cell line CMT-93 was a kind gift of Dr. D. Podolsky (Boston, USA). This cell line was cultured as previously described [15,32]. Human lung NCI-H292 cell line was cultured as previously described [33].

RT-PCR

Total RNAs from cultured cells were prepared using the QIAamp RNA blood mini-kit from Qiagen. Total RNA (1.5 μg) was used to prepare first strand cDNA (AdvantageTM RT-

1
2
3 for-PCR kit, BD Biosciences Clontech, Montigny-le-Bretonneux, France). PCR was
4
5 performed on 2 μ l of cDNA using specific pairs of primers as previously described [14]. The
6
7 annealing temperature was 58°C. *Muc5b* forward primer: 5'-
8
9 GAGGTCAACATCACCTTCTGC-3', *Muc5b* reverse primer: 5'-
10
11 TCTCATGGTCAGTTGTGCAGG-3'. β -actin was used as an internal control, mouse β -actin
12
13 forward primer: 5'-TCACGCCATCCTGCGTCTGGACT-3', mouse β -actin reverse primer:
14
15 5'-CCGGACTCATCGTACTCCT-3'. *Muc5b* [11] and β -actin PCR product sizes are 319 and
16
17 582 basepairs (bp), respectively. 100 bp DNA ladder was purchased from Amersham
18
19 Bioscience (Orsay, France). Densitometric analyses of PCR band for mMuc5b and β -actin
20
21 were performed using gel analyst software (Clara Vision, Paris, France).
22
23
24
25
26
27
28
29

30 *Transfections*

31
32 Transfection and co-transfection experiments were performed using Effectene[®] reagent
33
34 (Qiagen) as previously described [34]. Total cell extracts were prepared after a 48h incubation
35
36 at 37°C using 1X Reagent Lysis Buffer (Promega) as described in the manufacturer's
37
38 instruction manual. Luciferase activity (20 μ l) was measured on a Turner Design 20/20
39
40 luminometer (Promega). Total protein content in the extract (4 μ l) was measured using the
41
42 bicinchoninic acid method in 96-well plates as described in the manufacturer's instruction
43
44 manual (Perbio Sciences, Brebieres, France). Relative luciferase activity was expressed as
45
46 fold activation of luciferase activity by each deletion mutant compared with that of empty
47
48 pGL3 basic vector. In co-transfection experiments, 1 μ g of the deletion mutant of interest was
49
50 transfected with 0.25 μ g of the expression plasmid encoding the transcription factor of
51
52 interest. Results were expressed as fold activation of luciferase activity of the transcription
53
54 factor of interest compared to the co-transfection performed in the presence of the
55
56 corresponding empty control vector. Each plasmid was assayed in triplicate in three separate
57
58
59
60

1
2
3 experiments. To study the effect of transcription factor overexpression on endogenous *Muc5b*
4 mRNA level, cells (0.5×10^6) were transfected as before [15] with 4 μg of the expression
5 vectors of interest and cultured for 48h before being lysed and processed for total RNA
6 preparation and RT-PCR analysis. These experiments were performed in triplicate in three
7 independent series. The *Muc5b*/ β -actin ratio was calculated by densitometric analysis of the
8 DNA bands on the agarose gel by using the GelAnalyst-GelSmart software (Clara Vision).
9
10
11
12
13
14
15
16
17
18
19

20 ***Nuclear extract preparation***

21
22 Nuclear extracts from the CMT-93 and NCI-H292 cells, that express the different
23 transcription factors of interest, were prepared as described by Van Seuning *et al.* [46], and
24 kept at -80°C until use. Protein content (2 μl of cell extracts) was measured using the
25 bicinchoninic acid as described above.
26
27
28
29
30
31
32
33

34 ***Oligonucleotides and DNA probes***

35
36 The sequences of the oligonucleotides used for electrophoretic mobility shift assays (EMSA)
37 are indicated in table 2. They were synthesized by MWG-Biotech (Ebersberg, Germany).
38 Putative binding sites were identified using MatInspector (www.genomatix.de) and Match
39 and AliBaba 2.1 (www.gene-regulation.com) softwares. Consensus GATA probe was
40 purchased from SantaCruz laboratories. Equimolar amounts of single-stranded
41 oligonucleotides were annealed and radiolabeled using T4 polynucleotide kinase (Promega)
42 and [$\gamma^{32}\text{P}$]-dATP. Radiolabeled probes were purified by chromatography on a Bio-Gel P-6
43 column (Bio-Rad, Marnes-la-Coquette, France). The commercial GATA probe 5'-CAC TTG
44 ATA ACA GAA AGT GAT AAC TCT-3' was purchased from Santa Cruz Laboratory (sc-
45 2531).
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Electrophoretic mobility shift assay

EMSA were carried out as before [14]. Briefly, nuclear proteins (8 μg) were pre-incubated for 20 min on ice in 20 μl binding buffer with 1 μg of poly dI-dC (Sigma-Aldrich, Saint-Quentin Fallavier, France) and 1 μg sonicated salmon sperm DNA. Radiolabeled DNA probe was added (60,000 c.p.m.) and the reaction was left for another 20 min on ice. For super-shift analyses, 1 μl of the antibody of interest (anti-GATA-4, anti-GATA-5, anti-GATA-6, 0.2 mg/ml, SantaCruz Biotechnology, Tebu-Bio, Le Perray en Yvelines, France) was added to the proteins and left for 30 min at RT before adding the radiolabeled probe. Cold competition was performed by preincubating the nuclear proteins with a 50x excess of the unlabeled probe before adding the radioactive probe. Reactions were stopped by adding 2 μl of loading buffer. The GATA consensus probe was purchased from Santa Cruz Laboratory (sc-2531). Samples were loaded onto a 4% non-denaturing polyacrylamide gel and electrophoresis conditions were as described in [29]. Gels were vacuum-dried and autoradiographed overnight at -80°C .

Chromatin Immuno Precipitation (ChIP)

The Chromatin Immuno Precipitation (ChIP) assay was carried out as previously described [47] using 4 mg of anti-GATA-4, anti-GATA-5 (R&D) and anti-GATA-6 (N18 from Santa Cruz Biotechnology) antibodies or normal rabbit IgGs (Upstate, Millipore, St Quentin en Yvelines, France) with slight modifications. Immuno-precipitation was performed using Dynabeads® magnetic beads A and G (Invitrogen) with Dynabeads® rack (Invitrogen) following manufacturer's protocol. For the PCR, primers were designed to selectively amplify a -503/-261 region of *Muc5b* promoter. Forward primer 5'-CAG ACC CTC AGA AGC TAC A-3' and reverse primer 5'-CTA TGG GGT GGG TAT TTG-3'. PCR was carried out in a 30

1
2
3 µl volume containing 50 ng of DNA, 5U of AmpliTaq Gold (Applied Biosystems,
4 Courtaboeuf, France), 0.5 mM of each primer, 2.5mM MgCl₂ and 5% dimethylsulphoxide
5
6 using the following protocol: 3 min at 95°C followed by ((95°C) 15 s, (55°C) 15 s, (72°C) 15
7
8 s) for 34 cycles, and 72°C for 5 min. 242 bp PCR products were analysed on a 1.2 % (w/v)
9
10 agarose gel containing ethidium bromide.
11
12
13
14
15
16

17 *Immunohistochemistry*

18
19
20 Immunohistochemical studies for TTF-1 and MUC5B expression in lung
21 adenocarcinomas were performed as previously described [12]. TTF-1 monoclonal antibody
22 was purchased from DAKO (Trappes, France). MUC5B monoclonal antibody was provided
23 by Dr. D Swallow (MRC, London). The antibodies were used as followed : 1:2500 dilution of
24 goat anti-GATA-4 (R&D) or goat anti-GATA-6 (R&D) in PBS containing 1 % (w/v) bovine
25 serum albumin and 0.1% (v/v) Triton X-100. The sections were incubated for 1 h with
26 biotinylated horse anti-goat IgG (diluted 1:2000, Vector Laboratories, Biovalley, Marne la
27 Vallée, France). Sections were counterstained with haematoxylin, dehydrated and mounted. A
28 positive control for GATA-4 and GATA-6 immunostainings on human small intestine was
29 included in each set of experiments.
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

46 *Statistical analysis*

47
48 Statistical analyses were performed using Graphpad Prism 4.0 software. Data are
49 presented as mean±s.d. Differences in the mean of samples were analysed using ANOVA test
50 with selected comparison. P<0.05 was considered significant and indicated with an *. ***
51 indicates p<0.001
52
53
54
55
56
57
58
59
60

Acknowledgements

We thank Dr. M.-P. Buisine (Inserm U837, team 5) for the kind gift of Ali2 cosmid, Dr. S. Cereghini (Inserm U423, Paris, France) for the kind gift of the pMT2-GATA-4 expression vector, Dr. J. K. Divine (Washington University, St-Louis, MO, USA) for the kind gift of pSG5-GATA-5 and pSG5-GATA-6 expression vectors, and Dr. R. Di Lauro (Stazione Zoologica Anton Dohrn, Naples, Italy) for the kind gift of pCMV-TTF-1 expression vector. We are grateful to Dr. D. Podolsky (Massachusetts General Hospital, Boston, MA, USA) for providing us with the murine CMT-93 cell line. N. Jonckheere is the recipient of a Ligue Nationale contre le Cancer (LNCC) postdoctoral fellowship.

For Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- [1] Hollingsworth, MA and Swanson, BJ (2004) Mucins in cancer: protection and control of the cell surface, *Nat Rev Cancer* **4**, 45-60.
- [2] Dekker, J, Rossen, JW, Buller, HA and Einerhand, AW (2002) The MUC family: an obituary, *Trends Biochem Sci* **27**, 126-31.
- [3] Audie, JP, Janin, A, Porchet, N, Copin, MC, Gosselin, B and Aubert, JP (1993) Expression of human mucin genes in respiratory, digestive, and reproductive tracts ascertained by in situ hybridization, *J Histochem Cytochem* **41**, 1479-85.
- [4] Chen, Y, Zhao, YH, Di, YP and Wu, R (2001) Characterization of human mucin 5B gene expression in airway epithelium and the genomic clone of the amino-terminal and 5'-flanking region, *Am J Respir Cell Mol Biol* **25**, 542-53.
- [5] Copin, MC, Buisine, MP, Devisme, L, Leroy, X, Escande, F, Gosselin, B, Aubert, JP and Porchet, N (2001) Normal respiratory mucosa, precursor lesions and lung carcinomas: differential expression of human mucin genes, *Front Biosci* **6**, D1264-75.
- [6] Jonckheere, N and Van Seuningen, I (2010) The membrane-bound mucins: From cell signalling to transcriptional regulation and expression in epithelial cancers, *Biochimie* **92**, 1-11.
- [7] Rose, MC and Voynow, JA (2006) Respiratory tract mucin genes and mucin glycoproteins in health and disease, *Physiol Rev* **86**, 245-78.
- [8] Buisine, MP, Devisme, L, Copin, MC, Durand-Reville, M, Gosselin, B, Aubert, JP and Porchet, N (1999) Developmental mucin gene expression in the human respiratory tract, *Am J Respir Cell Mol Biol* **20**, 209-18.
- [9] Reid, CJ and Harris, A (1998) Developmental expression of mucin genes in the human gastrointestinal system, *Gut* **42**, 220-6.
- [10] Copin, MC, Devisme, L, Buisine, MP, Marquette, CH, Wurtz, A, Aubert, JP, Gosselin, B and Porchet, N (2000) From normal respiratory mucosa to epidermoid carcinoma: expression of human mucin genes, *Int J Cancer* **86**, 162-8.
- [11] Escande, F, Porchet, N, Aubert, JP and Buisine, MP (2002) The mouse Muc5b mucin gene: cDNA and genomic structures, chromosomal localization and expression, *Biochem J* **363**, 589-98.
- [12] Copin, MC, Buisine, MP, Leteurtre, E, Marquette, CH, Porte, H, Aubert, JP, Gosselin, B and Porchet, N (2001) Mucinous bronchioloalveolar carcinomas display a specific pattern of mucin gene expression among primary lung adenocarcinomas, *Hum Pathol* **32**, 274-81.
- [13] Van Seuningen, I, Pigny, P, Perrais, M, Porchet, N and Aubert, JP (2001) Transcriptional regulation of the 11p15 mucin genes. Towards new biological tools in human therapy, in inflammatory diseases and cancer?, *Front Biosci* **6**, D1216-34.
- [14] Mesquita, P, Jonckheere, N, Almeida, R, Ducourouble, MP, Serpa, J, Silva, E, Pigny, P, Silva, FS, Reis, C, Silberg, D, Van Seuningen, I and David, L (2003) Human MUC2 mucin gene is transcriptionally regulated by Cdx homeodomain proteins in gastrointestinal carcinoma cell lines, *J Biol Chem* **278**, 51549-56.
- [15] van der Sluis, M, Melis, MH, Jonckheere, N, Ducourouble, MP, Buller, HA, Renes, I, Einerhand, AW and Van Seuningen, I (2004) The murine Muc2 mucin gene is transcriptionally regulated by the zinc-finger GATA-4 transcription factor in intestinal cells, *Biochem Biophys Res Commun* **325**, 952-60.
- [16] Jonckheere, N, Vincent, A, Perrais, M, Ducourouble, MP, Male, AK, Aubert, JP, Pigny, P, Carraway, KL, Freund, JN, Renes, IB and Van Seuningen, I (2007) The human mucin MUC4 is transcriptionally regulated by caudal-related homeobox,

- 1
2
3 hepatocyte nuclear factors, forkhead box A, and GATA endodermal transcription
4 factors in epithelial cancer cells, *J Biol Chem* **282**, 22638-50.
- 5 [17] Cardoso, WV (1995) Transcription factors and pattern formation in the developing
6 lung, *Am J Physiol* **269**, L429-42.
- 7 [18] Costa, RH, Kalinichenko, VV and Lim, L (2001) Transcription factors in mouse lung
8 development and function, *Am J Physiol Lung Cell Mol Physiol* **280**, L823-38.
- 9 [19] Kimura, S, Hara, Y, Pineau, T, Fernandez-Salguero, P, Fox, CH, Ward, JM and
10 Gonzalez, FJ (1996) The T/eBP null mouse: thyroid-specific enhancer-binding protein
11 is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary,
12 *Genes Dev* **10**, 60-9.
- 13 [20] Maeda, Y, Dave, V and Whitsett, JA (2007) Transcriptional control of lung
14 morphogenesis, *Physiol Rev* **87**, 219-44.
- 15 [21] Liu, C, Glasser, SW, Wan, H and Whitsett, JA (2002) GATA-6 and thyroid
16 transcription factor-1 directly interact and regulate surfactant protein-C gene
17 expression, *J Biol Chem* **277**, 4519-25.
- 18 [22] Zhang, L, Whitsett, JA and Stripp, BR (1997) Regulation of Clara cell secretory
19 protein gene transcription by thyroid transcription factor-1, *Biochim Biophys Acta*
20 **1350**, 359-67.
- 21 [23] Ray, MK, Chen, CY, Schwartz, RJ and DeMayo, FJ (1996) Transcriptional regulation
22 of a mouse Clara cell-specific protein (mCC10) gene by the NKx transcription factor
23 family members thyroid transcription factor 1 and cardiac muscle-specific homeobox
24 protein (CSX), *Mol Cell Biol* **16**, 2056-64.
- 25 [24] Copin, MC, Perrais, M and Van Seuningen, I (2008) *Mucins in the pathophysiological*
26 *lung. Expression and roles in immuno- and gene-based therapies* in: *The Epithelial*
27 *Mucins: Structure/Function. Roles in Cancer and Inflammatory Diseases*, pp. 109-124
28 (Van Seuningen, I., Ed.) Research Signpost, Kerala, India.
- 29 [25] Molkenin, JD (2000) The zinc finger-containing transcription factors GATA-4, -5,
30 and -6. Ubiquitously expressed regulators of tissue-specific gene expression, *J Biol*
31 *Chem* **275**, 38949-52.
- 32 [26] Morrissey, EE, Ip, HS, Lu, MM and Parmacek, MS (1996) GATA-6: a zinc finger
33 transcription factor that is expressed in multiple cell lineages derived from lateral
34 mesoderm, *Dev Biol* **177**, 309-22.
- 35 [27] Morrissey, EE, Ip, HS, Tang, Z, Lu, MM and Parmacek, MS (1997) GATA-5: a
36 transcriptional activator expressed in a novel temporally and spatially-restricted
37 pattern during embryonic development, *Dev Biol* **183**, 21-36.
- 38 [28] Weidenfeld, J, Shu, W, Zhang, L, Millar, SE and Morrissey, EE (2002) The WNT7b
39 promoter is regulated by TTF-1, GATA6, and Foxa2 in lung epithelium, *J Biol Chem*
40 **277**, 21061-70.
- 41 [29] Van Seuningen, I, Perrais, M, Pigny, P, Porchet, N and Aubert, JP (2000) Sequence of
42 the 5'-flanking region and promoter activity of the human mucin gene MUC5B in
43 different phenotypes of colon cancer cells, *Biochem J* **348 Pt 3**, 675-86.
- 44 [30] Tan, D, Li, Q, Deeb, G, Ramnath, N, Slocum, HK, Brooks, J, Cheney, R, Wiseman, S,
45 Anderson, T and Loewen, G (2003) Thyroid transcription factor-1 expression
46 prevalence and its clinical implications in non-small cell lung cancer: a high-
47 throughput tissue microarray and immunohistochemistry study, *Hum Pathol* **34**, 597-
48 604.
- 49 [31] Yamazaki, K (2003) Pulmonary well-differentiated fetal adenocarcinoma expressing
50 lineage-specific transcription factors (TTF-1 and GATA-6) to respiratory epithelial
51
52
53
54
55
56
57
58
59
60

- 1
2
3 differentiation: an immunohistochemical and ultrastructural study, *Virchows Arch*
4 **442**, 393-9.
- 5
6 [32] Jonckheere, N, Van Der Sluis, M, Velghe, A, Buisine, MP, Suttmuller, M,
7 Ducourouble, MP, Pigny, P, Buller, HA, Aubert, JP, Einerhand, AW and Van
8 Seuningen, I (2004) Transcriptional activation of the murine Muc5ac mucin gene in
9 epithelial cancer cells by TGF-beta/Smad4 signalling pathway is potentiated by Sp1,
10 *Biochem J* **377**, 797-808.
- 11 [33] Perrais, M, Pigny, P, Copin, MC, Aubert, JP and Van Seuningen, I (2002) Induction of
12 MUC2 and MUC5AC mucins by factors of the epidermal growth factor (EGF) family
13 is mediated by EGF receptor/Ras/Raf/extracellular signal-regulated kinase cascade
14 and Sp1, *J Biol Chem* **277**, 32258-67.
- 15 [34] Perrais, M, Pigny, P, Buisine, MP, Porchet, N, Aubert, JP and Van Seuningen-
16 Lempire, I (2001) Aberrant expression of human mucin gene MUC5B in gastric
17 carcinoma and cancer cells. Identification and regulation of a distal promoter, *J Biol*
18 *Chem* **276**, 15386-96.
- 19 [35] Hackett, B, Bingel, C and Gitlin, J (1996) Mechanisms of gene expression and cell
20 fate determination in the developing pulmonary epithelium, *Annu Rev Physiol* **58**, 51-
21 71.
- 22 [36] Whitsett, JA (2002) Intrinsic and innate defenses in the lung: intersection of pathways
23 regulating lung morphogenesis, host defense, and repair, *J Clin Invest* **109**, 565-9.
- 24 [37] Nakamura, N, Miyagi, E, Murata, S, Kawaoi, A and Katoh, R (2002) Expression of
25 thyroid transcription factor-1 in normal and neoplastic lung tissues, *Mod Pathol* **15**,
26 1058-67.
- 27 [38] Hovenberg, HW, Davies, JR and Carlstedt, I (1996) Different mucins are produced by
28 the surface epithelium and the submucosa in human trachea: identification of
29 MUC5AC as a major mucin from the goblet cells, *Biochem J* **318 (Pt 1)**, 319-24.
- 30 [39] Zaret, K (1999) Developmental competence of the gut endoderm: genetic potentiation
31 by GATA and HNF3/fork head proteins, *Dev Biol* **209**, 1-10.
- 32 [40] Guo, M, Akiyama, Y, House, MG, Hooker, CM, Heath, E, Gabrielson, E, Yang, SC,
33 Han, Y, Baylin, SB, Herman, JG and Brock, MV (2004) Hypermethylation of the
34 GATA genes in lung cancer, *Clin Cancer Res* **10**, 7917-24.
- 35 [41] Capo-chichi, CD, Roland, IH, Vanderveer, L, Bao, R, Yamagata, T, Hirai, H, Cohen,
36 C, Hamilton, TC, Godwin, AK and Xu, XX (2003) Anomalous expression of
37 epithelial differentiation-determining GATA factors in ovarian tumorigenesis, *Cancer*
38 *Res* **63**, 4967-77.
- 39 [42] Ren, CY, Akiyama, Y, Miyake, S and Yuasa, Y (2004) Transcription factor GATA-5
40 selectively up-regulates mucin gene expression, *J Cancer Res Clin Oncol* **130**, 245-52.
- 41 [43] Buisine, MP, Porchet, N and Van Seuningen, I (2008) *Mucin expression and*
42 *regulation during development and cell differentiation* in: *The Epithelial Mucins:*
43 *Structure/Function. Roles in Cancer and Inflammatory Diseases*, pp. 75-94 (Van
44 Seuningen, I., Ed.) Research Signpost, Kerala, India.
- 45 [44] Naltner, A, Wert, S, Whitsett, JA and Yan, C (2000) Temporal/spatial expression of
46 nuclear receptor coactivators in the mouse lung, *Am J Physiol Lung Cell Mol Physiol*
47 **279**, L1066-74.
- 48 [45] Yang, H, Lu, MM, Zhang, L, Whitsett, JA and Morrissey, EE (2002) GATA6 regulates
49 differentiation of distal lung epithelium, *Development* **129**, 2233-46.
- 50 [46] Van Seuningen, I, Ostrowski, J, Bustelo, XR, Sleath, PR and Bomszyk, K (1995) The
51 K protein domain that recruits the interleukin 1-responsive K protein kinase lies
52
53
54
55
56
57
58
59
60

1
2
3 adjacent to a cluster of c-Src and Vav SH3-binding sites. Implications that K protein
4 acts as a docking platform, *J Biol Chem* **270**, 26976-85.

- 5 [47] Piessen, G, Jonckheere, N, Vincent, A, Hemon, B, Ducourouble, MP, Copin, MC,
6 Mariette, C and Van Seuning, I (2007) Regulation of the human mucin MUC4 by
7 taurodeoxycholic and taurochenodeoxycholic bile acids in oesophageal cancer cells is
8 mediated by hepatocyte nuclear factor 1alpha, *Biochem J* **402**, 81-91.
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

Figure legends

Figure 1: Sequence of the promoter of murine *Muc5b* mucin gene. (A): *Muc5b* promoter is flanked by a TATA box (double-underlined). The broken arrow indicates the position of the transcription start site designated +1. The first ATG is bold and italicized. Gray boxes indicate putative binding sites for transcription factors and boxed sequences indicate the sequences of oligonucleotides used in EMSA. Identified transcription factors by EMSA are in bold. (B): Alignment of the proximal part of mouse *Muc5b* and human *MUC5B* promoters. Conserved nucleotides are in gray and conserved TATA box is bold and boxed.

Figure 2: Characterization of *Muc5b* promoter activity in CMT-93 and NCI-H292 cancer cell lines by transient transfection. (A): Expression of *Muc5b* by RT-PCR in CMT-93 cells. 2 and 10 μ l of β -actin (lane 2) and *Muc5b* (lane 3) PCR products were respectively loaded on a 1.5% agarose gel containing ethidium bromide. 100 bp ladder (lane 1). (B): Schematic representation of the different deletion mutants used to study *Muc5b* promoter activity. Numbering refers to transcription initiation site designated as +1, (C): Luciferase activity diagram showing *Muc5b* promoter activity in CMT-93 (black bars) and in NCI-H292 (gray bars) cells. 1 μ g of each pGL3-*Muc5b* deletion mutant was transfected as described in the materials and methods section. Results are expressed as fold activation of luciferase activity of the deletion mutant of interest compared to the activity of empty pGL3 basic vector (white bar). Standard deviation represents the means of values obtained in triplicate in three separate experiments.

Figure 3: Regulation of *Muc5b* promoter by TTF-1 transcription factor. Identification of TTF-1 cis-elements by EMSA. (A): Measurement of *Muc5b* mRNA level by RT-PCR in

1
2
3 CMT-93 cells transfected with either 4 μg of pCMV-TTF-1 (TTF-1) or 4 μg pCMV4 empty
4 vector (Ref.). The diagram represents the calculated ratio of *Muc5b*/ β -actin. Standard
5 deviation represents the means of values obtained from three separate experiments. (B): Co-
6 transfection experiments in CMT-93 (black bars) and in NCI-H292 (gray bars) cells were
7 performed in the presence of 1 μg of *Muc5b* pGL3-deletion mutants and 0.25 μg of pCMV-
8 TTF-1 expression vector. Ref. refers to the normalized luciferase activity of the pGL3-
9 deletion mutants of interest co-transfected with the empty expression vector pCMV4.
10 Luciferase activity for each co-transfection is represented as fold activation compared to
11 activity obtained with empty pCMV4 vector. Standard deviation represents the means of
12 values obtained in triplicate in three separate experiments. (C): Identification of TTF-1 *cis*-
13 elements by EMSA. 8 μg of nuclear extracts from NCI-H292 cells were incubated with the
14 radiolabeled DNA probes as indicated. Lanes 1-4, T211, TTF-1 sites at -358/-355 and -353/-
15 35; lanes 5-6, mutated T211; lanes 7-10, T212, TTF-1 sites at -709/-706 and -700/-697; lanes
16 11-12, mutated T212. Radiolabeled probe alone (lanes 1, 5, 7, and 11). Incubation of T211,
17 mut. T211, T212, or mut. T212, probes with NCI-H292 nuclear proteins (lanes 2, 6, 8, and
18 12). Cold competition with 50x excess of cold T211 (lane 3), mutated cold T211 (lane 4),
19 cold T212 (lane 9) or mutated cold T212 (lane 10) probes, respectively. DNA-protein
20 complexes (TTF-1) are indicated by an arrow on both sides of the autoradiograms. The
21 asterisk highlights the TTF-1 band decreased by T212 probe cold competition (lane 9).
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 **Figure 4: Regulation of *Muc5b* promoter by GATA-4/-5/-6 transcription factors.**

52 **Identification of a GATA *cis*-element by EMSA.** (A): Measurement of *Muc5b* mRNA level
53 by RT-PCR in CMT-93 cells transfected with either 4 μg of pMT2-GATA-4 (GATA-4),
54 pSG5-GATA-5 (GATA-5), pSG5-GATA-6 (GATA-6) or 4 μg of the corresponding empty
55 vector (Ref.). The diagram represents the calculated ratio of *Muc5b*/ β -actin. Standard
56
57
58
59
60

1
2
3 deviation represents the means of values obtained from three separate experiments. **(B)**: Co-
4 transfection experiments in CMT-93 cells were performed in the presence of 1 μg of *Muc5b*
5 pGL3-deletion mutants and 0.25 μg of pMT2-GATA-4 (white bars), pSG5-GATA-5 (black
6 bars) or pSG5-GATA-6 (gray bars) expression vectors. Ref. refers to the normalized
7 luciferase activity of the pGL3-deletion mutants of interest co-transfected with the
8 corresponding empty vectors, respectively. Luciferase activity for each co-transfection is
9 represented as fold activation compared to activity obtained with empty vector. Standard
10 deviation represents the means of values obtained in triplicate in three separate experiments.
11 Statistical analysis was performed using ANOVA test with selected comparisons. * indicates
12 $p < 0.05$. *** indicates $p < 0.001$. **(C)**: Co-transfection experiments in NCI-H292 cells were
13 performed under the same conditions as in CMT-93 cells. **(D)**: Identification of GATA *cis*-
14 elements by EMSA. 8 μg of nuclear extracts from CMT-93 cells were incubated with the
15 T242 (GATA at -411/-408), T254 (GATA at -454/-449) and T84 (GATA at -1143/-1140)
16 radiolabeled DNA probes. Radiolabeled probes alone (lanes 1, 8 and 14), incubation of T242,
17 T254 and T84 probes with CMT-93 nuclear proteins (lanes 2, 9 and 15), cold competition
18 with 50x excess of cold T242 (lane 3), mutated cold T242 (lane 4), cold T254 (lane 10), and
19 cold T84 (lane 16), supershift analysis with anti-GATA-4 (lanes 5, 11 and 17), anti-GATA-5
20 (lanes 6, 12 and 18), and anti-GATA-6 (lane 7, 13 and 19) antibodies, respectively. DNA-
21 protein complex (GATA) and supershift (ss GATA-4, ss GATA-6) are indicated by an arrow
22 on both sides of the autoradiograms. **(E)** In vivo binding of GATA-4, -5, and -6 to chromatin
23 by ChIP in CMT-93 cells. PCRs were carried out with specific pairs of primers covering
24 GATA sites. PCR products (10 μl) were analyzed on 1.2 % (w/v) agarose gels. IgGs, negative
25 control with rabbit IgGs. **(F)**: Study of synergistic activity between TTF-1 and GATA-6 on
26 *Muc5b* promoter. Co-transfection experiments were performed in CMT-93 cells in the
27 presence of 1 μg of *Muc5b* pGL3-deletion mutants as indicated and 0.25 μg of pCMV-TTF-1,
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 0.25 μ g of pSG5-GATA-6 or both. Results are expressed as fold activation of luciferase
4
5 activity obtained in co-transfected with expression vector encoding the transcription factor of
6
7 interest or both compared with cells transfected with corresponding empty vectors (Ref.).
8
9 Standard deviation represents the means of values obtained in triplicate in three separate
10
11 experiments.
12
13
14
15
16

17 **Figure 5: Expression of MUC5B, TTF-1 and GATA-6 in several types of human lung**
18 **adenocarcinomas.** Immunohistochemistry was performed as described in the materials and
19
20 methods. Well-differentiated lung adenocarcinoma stained strongly for MUC5B (A), not for
21
22 TTF-1 (B) and stained for GATA-6 (C), x100 magnification. Papillary lung adenocarcinoma
23
24 stained for MUC5B (D), TTF-1 (E) and GATA-6 (F), x200 magnification. Focal mucinous
25
26 area of a non mucinous bronchioalveolar carcinoma stained for MUC5B (G) and not for TTF-
27
28 1 (*inset*), x400 magnification. (H): Same tumor as in (G), non mucinous bronchioalveolar
29
30 carcinoma stained for TTF-1 and not for MUC5B (*inset*) and stained for GATA-6 (I), x200
31
32 magnification.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Position in the promoter	Oligonucleotide sequences (5'→3')	Orientation
<i>Muc5b</i>		
-1195/-1	CGC <i>GAG CTC</i> CAC ATA GAC TTT TCC CTT	S
	CGC <i>ACG CGT</i> GGC ACA GTG ATG TAA ATC	AS
-717/-1	CGC <i>GAG CTC</i> CCA GGG CCC TTG AGA C	S
	CGC <i>ACG CGT</i> GGC ACA GTG ATG TAA ATC	AS
-478/-1	CGC <i>GAG CTC</i> CAG GGA CCC TGC CAG	S
	CGC <i>ACG CGT</i> GGC ACA GTG ATG TAA ATC	AS
-169/-1	CGC <i>GAG CTC</i> TTG CTC CCT GGG GGC CTG	S
	CGC <i>ACG CGT</i> GGC ACA GTG ATG TAA ATC	AS
-717/+47	CGC <i>GAG CTC</i> CCA GGG CCC TTG AGA C	S
	CGC <i>ACG CGT</i> CCT GGG GGC AGT ACA	AS
-478/+47	CGC <i>GAG CTC</i> CAG GGA CCC TGC CAG	S
	CGC <i>ACG CGT</i> CCT GGG GGC AGT ACA	AS

Table 1: Sequences of the pairs of oligonucleotides used in PCR to produce deletion mutants covering murine *Muc5b* promoter. *Sac*I (GAGCTC) and *Mlu*I (ACGCGT) sites (bold and italicized) are added at the end of the primers to direct subcloning into pGL3 basic vector. S: sense, AS: antisense.

Probe	Putative binding site	Sequence (5' → 3')
	<i>Muc5b</i>	
T213	TTF-1 (-112/-109)	CTGCCATGGCCCCTCCC <u>CAAGAG</u> CAAA
T211	TTF-1 (-358/-355; -353/-350)	CGGCAAACACA <u>AGCCAAGG</u> TTGTTGTC
T240	Mutated T211	CGGCAAACAG <u>TATCGTAT</u> GTTGTTGTC
T212	TTF-1 (-709/-706; -700/-697)	TCCAGGGCC <u>CTTGAGACCCTT</u> GGTTCATTTC
T241	Mutated T212	TCCAGGGCC <u>AGTAAGACCAGTA</u> GTCATTTC
T238	TTF-1 (-325/-322)	CCCCTGATC <u>CTTG</u> TAGTGTCTAGT
T84	GATA (-1143/-1140)	TCTCAGAAAG <u>GATA</u> AAGGATGGGGGC
T242	TTF-1/GATA (-417/-414; -411/-408)	TCACAGC <u>CTTGTTGATA</u> CTTTGGGGAC
T239	Mutated T242	TCACAGCCTTGTT <u>CTTACT</u> TTGGGGAC
T254	GATA (-454/-449)	GCCCATGAC <u>CATCTG</u> GAGCATAAT
consensus GATA	GATA	CAC TTG <u>ATA</u> ACA GAA AGT <u>GATA</u> AAC TCT

Table 2: Sequences of the sense oligonucleotides used for EMSAs. Antisense oligonucleotides were also synthesized and annealed to the sense oligonucleotides to produce double-stranded DNA. Positions of the putative binding sites are italicized and underlined. Mutated bases are bold and underlined.

A

-1210 AGTTTACAGACTTTTTCACATAGACTTTTCCCTTTCCTGTTCCATCTCTAGGAATTTAG **TTC**

GATA-4

-1150 **TCAGAAAAGATAAGGATGGGGGC**AGACAGGGAGGGGTTTTTCGTGATACCAAAGTCACTGA

-1090 ACACTGTGCCACTCCCTGACATGGCCAACCCCTAGGAACTTTCAGGGGTGGTCAGATCTG

-1030 GTGTCCATTCCTGGTGGCTATAACAGAAGGAATGTGACTGTATTCTGTATGTATCTAACA

-970 GGGAAAGAGAAACTTAGACGTGTGCACCCTGCTCTTTGTGGCAGTGTAGAAACGGCGTTGG

-910 TGGATATGTTTCATGGTTCAGGAAATGGCTCCCCAGAACCCCTCAGGAAACCCAACCAACA

-850 AATTAAGAGAGCTGAGCCCCCGGGACTCCCAATTCAGGATGGACAGTGGGTCAGGAGACT

-790 CCCGGTTATGCCCTTGCCCCACTGTCCCGCAGTTTGGGGTTTACAAGTTCCAGAGCTGAG

TTF-1 **TTF-1**

-730 GCCTGAGGTAAG**TCCAGGGCCCTTGAGACCCTTGGTCAATTTT**C TTTGTCAAGGCCCAGGG

-670 TTAATAAGCTGACCCTGGGAAGAGGTGGAGTCCAACCTAGAGAAGCTCATGGCCTGCCTTG

-610 GCCAAAACCAAGTACTGGGCCTCATTTTCTCCATGTGCGTGAGCTCTACAGCAGACAGGAG

-550 CCCAGAGACTCATCCTTTCTCAGATTCTGTGAGGGTTCCCTTCTCCAAGACCCCTCAGA

GATA-4/-6

-490 AGCTACACAGAACAGGGACCCCTGCCAG**GCCCCATGACCCTCTGGAGCATAAT**TTTCTGAGC

TTF-1 **GATA-4**

-430 CCACC**TCACAGCCTTGTGATACTTTGGGGAC**CCAGTGCCTAGCACCTGGGGTCAGCTT

TTF-1 **TTF-1** **TTF-1**

-370 CC**CGGC AAAACA CAAGCC AAGGTTGT**CTCTT**GCCCCGATCCCTTG**TAGTGTCTAGT

Sp1

-310 CTCTTAGAGGAGCTGACACAGGGCCCGCTGGCCAAATACCCACCCCATAGGCCCAGGGCT

-250 TCGAATCCAGAGTAGTCATCTCCCTACCCTGTCCA CTCCCTCCAGTTCTCAGAAGCCAC

-190 AAGGGGTGGGAGCACTCTGGTGGCACAGTGATGTAATCCTTCTTCTCCATGGCCTCTT

TTF1

-130 **CTGCCATGGCCCTCCCAAGAGCAAA**CACACGTGGCTGTGCAGGGAGTAGACCGTGCC

-70 TTACATGGCCAGGGA ACTGGGCTCTGGGCCAAAGCCTTTGGC**TACATAA**GTCCAGGCCCC

+1

-10 CAGGGAGCAAAGTGCCTTGTCTCAGTCCCTCCTGCCCTGCCCTCTGTACTGCCCCAGGATG

B

Muc5b GGCACAGT GATGT AAATCCTTCCTTCTCCATGGCCTCTT CCTGCCA TGGCCCTCCCCA A -110

MUC5B GGCACAGA GCTGCAAAATCCTTCCTGATCCAAGGCCTCTCC CCTGCCA CAGCCCTCCCCGA

Muc5b GAGCAAACACACGTTGGCTGTGCAGGGAGTAGACCCTTACATGGCCAGGGAAC TGGG -50

MUC5B GAGCAAACACACGTTGGCTGAGAGCGGGGAAGA GCA CGGTGCC CTGCGTGGCCTGGCCTGG

Muc5b CTCTGGGCCAAA GCCTTTGGC **TACATAA**

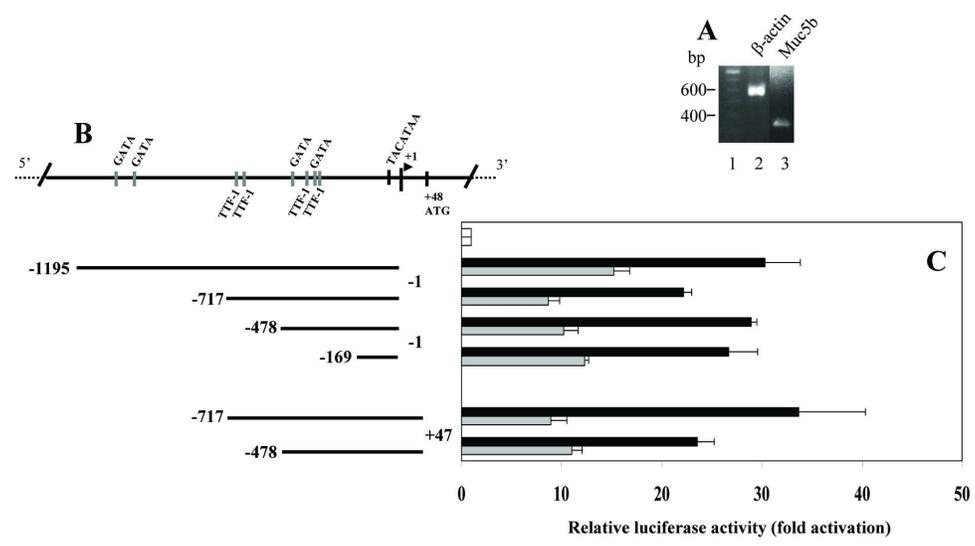
MUC5B CT TGGGGCAAAGGCTCCCTGC **TACATAA** -22

Figure 1 : Jonckheere *et al.*

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

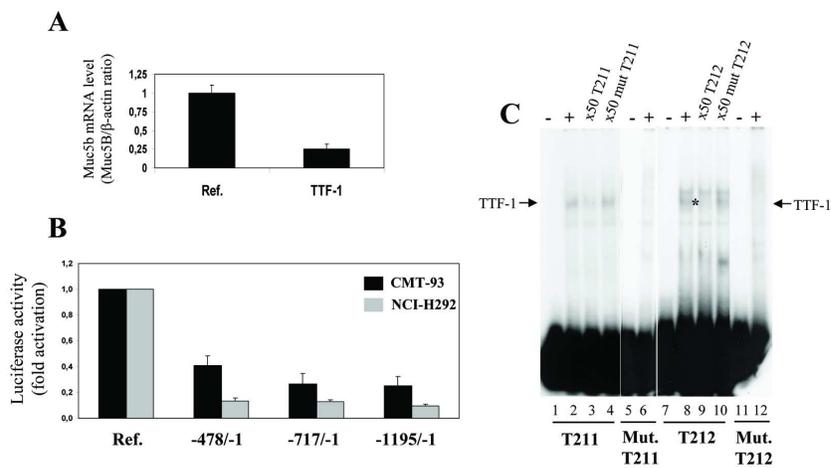
Figure 2 : Jonckheere *et al.*



165x124mm (300 x 300 DPI)

View Only

Figure 3 : Jonckheere *et al.*

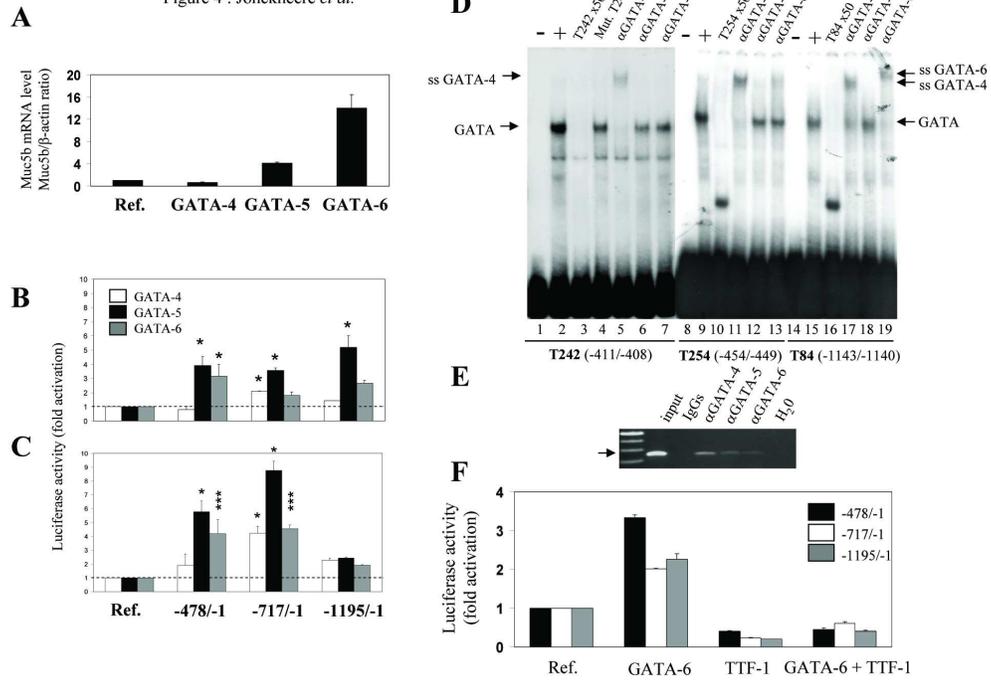


165x124mm (300 x 300 DPI)

Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

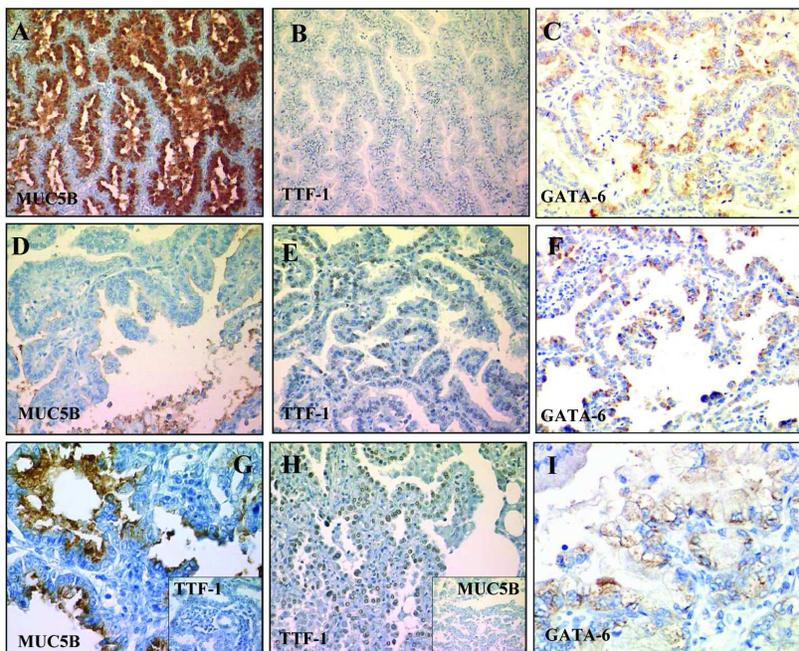
Figure 4 : Jonckheere *et al.*



165x124mm (300 x 300 DPI)

View Only

Figure 5 : Jonckheere *et al.*



165x124mm (300 x 300 DPI)

Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60