Induced sputum: a window to lung pathology

Ben Nicholas¹ and Ratko Djukanović
Division of Infection, Inflammation and Immunity, University of Southampton School of Medicine, Tremona Road, Southampton SO16 6YD, U.K.

Abstract
Sputum is recognized as a sampling method for the monitoring and assessment of chronic lung diseases such as asthma, COPD (chronic obstructive pulmonary disease) and cystic fibrosis. Sputum samples the central airways and its protein components (e.g. mucins and cytokines), cellular components (e.g. eosinophils and neutrophils) and microbiological components (e.g. viruses and bacteria) can be used as markers of disease severity, exacerbation, susceptibility or progression. This paper describes the basic constituents of induced sputum and how these influence the quantification and identification of novel biomarkers of chronic lung diseases using techniques such as proteomics.

Clinical utility of sputum
Induced sputum is recognized as a very useful sampling method for both research and clinical use aiding both the diagnosis and monitoring of disease status, particularly in relation to chronic lung diseases such as asthma, COPD (chronic obstructive pulmonary disease) and, increasingly, interstitial lung disease [1,2]. A great advantage of the technique is that it enables sampling of the airways in a non-invasive manner [3], in contrast with other methods such as bronchial biopsy, bronchial brushing and BAL (bronchoalveolar lavage), all of which require bronchoscopy and the discomfort and risk that it entails. This is particularly important in the examination of patients with severe airways disease where endoscopy poses significant risk to oxygen saturation levels. While sputum induction is not completely without risk in such patients [4], steps such as concomitant application of bronchodilators and/or stepwise increase in saline application can serve to reduce risk of the procedure [5].

A further advantage of induced sputum as a sampling method is the large body of knowledge now available on its characteristics in both normal subjects and patients with disease, particularly in relation to its inflammatory cell content, but also relating to other physical properties such as purulence and bacterial load [6,7].

A greater future potential for induced sputum has also been recognized within the last 10 years: since it samples the airways, sputum may contain protein/peptide components that could act as biomarkers of disease presence or severity. One caveat, however, is that, despite well-defined protocols for its collection and processing, sputum sampling is known to be variable, although stable enough for repeat samplings [2], and its viscosity requires further steps in processing, which may add to this variability. The present paper describes the current state of knowledge regarding the utility of induced sputum for biomarker discovery and quantification in the examination of pulmonary disease.

Collection of induced sputum
The method of sputum sampling to some extent influences its final composition, since the balance of tracheobronchial secretions to other biofluids may be altered through contamination. Thus it is generally accepted that, through travel via the buccal cavity to the sample dish, sputum will be contaminated by saliva, and if its collection requires significant expectoration (often the case in healthy non-smoking volunteers), may also be contaminated by plasma proteins through increased exudate induced by the expectoration process. Therefore one might expect differences between subjects with healthy or irritated airways independent of any disease process, simply reflecting differences in the ease of expectoration [8]. In addition, repeated sputum sampling can itself influence sputum composition through induction of a localized inflammatory response [9]. Therefore, a standardized protocol of sputum induction applied to all patients, using nebulized saline to loosen phlegm and aid expectoration, is critical for sputum sampling, especially where comparisons in biomarkers of disease are to be made.

Composition of sputum
Contaminants apart, sputum consists of constituents derived from the respiratory epithelium, mainly sampling the central airways [10]. Additional components are provided by transepithelial exudate from tissue plasma, salt, lipids, inflammatory cell components from cells resident within either the tissue or the airways lumen, externally derived particulate/inhaled matter from the environment, microbial products from any colonizing bacteria or viruses, and also cellular debris from all of these compartments.

Thus it is composed of a mixture of mucins (approx. 2–4% of the final weight) and other exudates with physiological relevance to the airways. The gel forming mucins MUC5B and MUC5AC are the predominant mucin forms in the airways [11] and are secreted in the submucosal glands of the epithelium and the goblet cells of the epithelium respectively.

Mucins are large glycoproteins made up of tandemly repeating amino acids (serine- and threonine-rich), which are O-linked glycosylation sites [12]. The high carbohydrate

Key words: biomarker, chronic lung disease, mucin, proteomics, sputum
Abbreviations used: BAL, bronchoalveolar lavage; BALF, BAL fluid; COPD, chronic obstructive pulmonary disease; DTT, dithiothreitol.
¹To whom correspondence should be addressed (email bln1@soton.ac.uk).
content of the mucin structure is thought to stiffen the protein filaments, increasing the volume domain of the molecule and accounting for some of its gel forming properties. Many bacteria also bind to certain sugar moieties, and thus bacterial binding and clearance may be the primary function of mucins [11]. The mucin monomers or subunits are approx. 2–3 MDa, but are capable of forming an oligomerized meshwork through a mixture of different molecular interactions (including hydrogen bonds, van der Waals forces and ionic bonds), which tangles to form a low-viscosity network [13]; furthermore, disulfide linkages can result in polymer gels of much greater size. The gel forming mucins provide a protective coating to the airways epithelium. They are diverse in molecular mass, ranging from 5 to 50 MDa [14], and appear filamentous when viewed using electron microscopy.

The normal movements of sputum in the mucociliary escalator ensure that there is constant movement of these components to prevent epithelial cell damage from its components. Effective clearance of mucus requires a balance of sputum volume and viscoelasticity along with an effective ciliary apparatus and pericilliary liquid volume.

Mucins are responsible for most of the physical characteristics of mucus, conferring elasticity and other cohesive/ adhesive properties. Mucus is a gel that acts in a non-Newtonian liquid manner (i.e. has liquid and solid behaviour) such that in normal secretions it can enable cilia to convert kinetic energy into mucus movement. In models, it has been shown that increasing viscosity will improve mucus clearance [15].

A change in sputum volume, of necessity accompanied by changes in physical characteristics, is a feature of airways disease. One consequence of airway epithelial damage is replacement of ciliated cells by goblet cells, a process that is largely under the control of the EGFR (epidermal growth factor receptor) signalling cascade [16]. This is thought to occur as a physiological response to increase airways irritation, resulting in improved clearance of pathogens and irritants, and demonstrates the dynamic nature of mucus control, alongside the physical barrier element of mucus.

Where sputum volume increases sufficiently, or where its physical properties are altered, preventing adequate clearance, resultant coughing occurs to aid clearance of the airways [17]. Other constituents of mucus are known to affect these properties. For example, there are significant quantities of DNA and actin present in the mucus of patients with various airways diseases derived from necrotic inflammatory cells in the lumen, and these can act to reduce elasticity or prevent clearance by changing sputum physical properties [18]. In addition to changes in the total quantity of mucin forms, changes in the relative proportions of their glycosylated isoforms may well contribute to airway physiology through knock-on effects on gel forming properties [19].

**Mucins and protein binding effects**

As previously indicated, the primary role of mucus appears to be an innate defence mechanism by the epithelium. Mucus forms a protective barrier, trapping airways particulates and transporting them to the external environment. A further biological effect of mucins occurs through their unique biochemical properties. Mucins are capable of binding a number of molecules in biofluids, and recent work categorizing the proteome of airway epithelia has identified proteins that bind with particular affinity to mucins, inferring potential importance in disease progression [20]. Such molecules include secretory IgA and its transporter molecule in the epithelium, the polyIg receptor, proline-rich proteins, transferrins, lysozyme and defensins [11]. Such molecules are part of the structural cell innate immunity to microbial infections and, in the case of IgA, provide a link between the acquired and innate immune cell defences also. Other proteins unique to induced sputum mucin-rich fractions, and not present in apical secretions of differentiated epithelial cell secretions, are in the majority derived from innate immune cells such as neutrophils [20].

A number of natural modulators of mucous gels have been found. These include substances that act as mucolytics, including gelsolin, a substance that reduces F-actin (filamentous actin) to its soluble form [18], thioredoxin [21] and bacterial chitinases [22], all natural secretory products of either healthy or diseased airways. Other substances can act upstream to prevent epithelial cell damage in response to oxidant stress such as smoking and consequent goblet cell hyperplasia, such as the antioxidant haem oxygenase 1 [23].

**Factors affecting mediator measurement in sputum**

As previously mentioned, sputum is a complex biofluid and has several properties which make measurement of its components more difficult: (i) The non-mucin protein component is only a small proportion of the total sputum contents. (ii) Liquefaction agents such as DTT (dithiothreitol) or NAC (N-acetylcysteine) are chemically active and may modify the components such that they are no longer measurable [24]. (iii) Akin to BAL, there is no endogenous component useful for normalization of quantitative data. (iv) Sputum is a salty biofluid and this ionic component can adversely affect biochemical analyses (e.g. isoelectric focusing, see Figure 1). (v) Sputum is prone to contamination by plasma, cells, saliva, microbes etc., all of which make the interpretation of results more challenging.

To overcome these difficulties, several modifications to the collection protocol can be made, including: selection of ‘mucoid’ components, which reduces salivary contamination, centrifugation or filtration to remove cell/microbial products, renaturation of sputum by dialysis to refold proteins and reduce salt content after DTT treatment [25] and weighing of selected sputum to normalize for sputum weight. Without such modifications, quantification of biomarkers in sputum has seen a great deal of variability or significant numbers of samples with zero data values. In an effort to redress this issue, several criteria for quantification of markers in sputum have been proposed, including the use of sample ‘spikes’ with recombinant proteins to determine percentage recoveries of...
Figure 1 | High salt and mucin content adversely affects isoelectric focusing of DTE (dithioerythritol)-processed sputum samples, resulting in poor resolution

Removal of these contaminants by size-exclusion filtration and subsequent dialysis restores spot resolution on two-dimensional gels (B. Nicholas, unpublished work).

each analyte as a measure of the effect of changes in sputum quality in disease on the marker quantification itself [24,26].

Quantification of markers of airways disease in sputum samples

A number of publications exist that describe quantification of markers in sputum and their change with disease. These studies have varied widely in their processing, normalization and assay validation protocols. Altered levels of a number of markers in sputum have been observed in stable COPD, including those related to remodelling [27], chemoattraction of inflammatory cells related to disease pathology [28,29] and co-morbidities such as metabolic dysfunction reflected in leptin levels [30]. Changes have also been observed during exacerbations of COPD [31,32] and in comparisons of COPD with other airways diseases such as cystic fibrosis [33]. However, a disappointing element of studies using known biomarkers is that a number of candidate biomarkers appear to be universally altered in airsweeps, irrespective of disease aetiology. These studies are also limited by the literature on biomarkers, which may have a bias towards systemic rather than locally produced biomarkers due to a historical reliance on blood-derived samples for analyses. Therefore there has been increasing interest in identifying novel biomarkers of disease, and sputum, through its non-invasiveness and direct sampling of the airways, provides an excellent potential sample in which to find such biomarkers.

Proteomics of induced sputum

There have been relatively few proteomic studies of induced sputum. It would be fair to say that cystic fibrosis has occupied the highest profile in research using sputum, both in terms of bacteriology and mucolytic therapy; however, due to the recognized mechanisms of this disease, there is possibly a lower motivation for performing expensive and time-consuming unbiased proteomic studies for surrogate markers for the disease. A few studies have, nevertheless, been performed, looking at biomarkers in BAL [34] and sputum [35] with the aim of understanding cystic fibrosis disease pathology, and many have examined the change in bacterial phenotype with colonization and antibiotic resistance [36,37]. The most interesting evidence in disease, however, lies in the reduction of intact mucins in the sputum of cystic fibrosis patients, which belies the common belief that cystic fibrosis results in a straightforward increase in sputum volume, which is detrimental to cough clearance. The decrease in intact mucin in cystic fibrosis sputum may be due to increased degradation and dilution with other cellular products [38]. Such information has been discovered by altering sputum analysis methods to include the high-molecular-mass material, which would normally be excluded using gel-based fractionation methods [20].

An increasing number of proteomic studies are investigating the effects of smoking and chronic disease in the pulmonary compartment. Benchmark proteomic studies in the form of so-called ‘shotgun’ analyses have been used to generate proteome lists for many compartments of the airways, including a number of biofluids, including saliva [39,40], BALF (BAL fluid) [41,42], nasal lavage fluid/mucus [43,44], epithelial lining fluid [45] and sputum [46]. Such studies are beginning to enable comparisons of proteomes between different airways compartments and different diseases (Figure 2).

A wide variety of unbiased proteomic studies comparing health with disease in the pulmonary compartment have identified mechanisms of disease or surrogate biomarkers of disease [47]. Comparative unbiased studies of sputum involving large numbers of patients or independent cohorts have not been reported; however, pilot studies have identified candidate biomarkers for diseases such as COPD [48], cystic fibrosis [35] and bronchiolitis [49]. Interesting data are emerging from these types of studies. In common with proteomic studies from other mucosal sites, the proteomic signature from the pulmonary compartment in response to any disease/stimulus appears to be made up of several categories of proteins, including those arising from the systemic compartment, probably present through increased tissue leakage such as albumin, immunoglobulins and α1-antitrypsin, those of the innate structural cell defences such as PLUNC (palate, lung and...
nasal epithelium carcinoma-associated protein), and those of the innate immune defence system, such as neutrophil- and macrophage-derived proteins, including lipocalin-2 and myeloperoxidase, and other epithelial cell-derived secretory products, which may be the more ideal candidates for novel disease biomarkers. One cannot forget also that local modulation of systemic and locally produced proteins through, for example, proteolytic degradation, may account for some of the biomarkers found using proteomic studies in biofluids, although often the modulator itself remains cryptic.

In conclusion, sputum is of increasing importance as a diagnostic tool for the diagnosis and monitoring of disease, the study of disease pathogenesis and its treatment, through its inflammatory cell, bacterial, volatiles, mucin and protein content. Measurement of these components is increasingly sophisticated and quantifiable, and the search for novel biomarkers of airways disease in this biofluid is under way. However, proteomic analyses to date have not provided any clinically proven biomarkers of lung disease. With the rapid pace of development of MS methods, with improved accuracy and throughput, and avoiding traditional gel-based applications, sputum will be increasingly utilized for such proteomic studies.

**Funding**

Work is supported by the National Institutes of Health [grant number RO1-HL72356-01].

**References**


©The Authors Journal compilation ©2009 Biochemical Society


27 Çalıkoglu, M., Unlu, A., Tamer, L. and Özgu, E. (2006) MMP-9 and TIMP-1 levels in the sputum of patients with chronic obstructive pulmonary disease and asthma. Tuberk. Toraks 54, 114-121


31 Xia, W., Hsu, Y.P., Ishizaka, A., Kirikae, T. and Moss, R.B. (2005) Sputum cathelicidin, urokinase plasminogen activation system components, and cytokines discriminate cystic fibrosis, COPD, and asthma inflammation. Chest 128, 2316-2326


Received 4 March 2009 

doi:10.1042/BST0370868