Label free electrical detection of DNA hybridisation for the example of influenza virus gene sequences

Andreas Kukol[1§], Peng Li[2], Pedro Estrela[2], Paul Ko-Ferrigno[3,4], Piero Migliorato[2]

[1]School of Life Sciences, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK
[2]Centre for Advanced Photonics and Electronics, Electrical Engineering Division, Department of Engineering, University of Cambridge, 9 JJ Thomson Avenue, Cambridge CB3 0FA, UK
[3]Hutchison/MRC Research Centre and Cancer Cell Unit, Hills Road, Cambridge CB2 0XZ, UK
[4]Present address: Leeds Institute of Molecular Medicine, St. James University Hospital, Leeds LS9 7TF, UK

[§]Corresponding author: Tel.: 01707 284543

Email addresses:
AK: a.kukol@herts.ac.uk
PL: pl281@cam.ac.uk
PE: pme22@cam.ac.uk
PKF: p.koferrigno@leeds.ac.uk
PM: pm128@cam.ac.uk

Short title: Label free electrical detection of hybridisation

Category: DNA recombinant techniques and nucleic acids

Abstract

Microarrays based on DNA-DNA hybridisation are potentially useful for detecting and subtyping viruses but require fluorescence labelling and imaging equipment. We investigate a label free electrical detection system using electrochemical impedance spectroscopy that is able to detect hybridisation of DNA target sequences derived from avian H5N1 influenza virus to gold surface-attached single stranded DNA oligonucleotide probes. A 23nt probe is able to detect a 120nt base fragment of the influenza A hemagglutinin gene sequence. A novel method of data analysis is described that is compatible with automatic measurement without operator input contrary to curve fitting used in conventional EIS data analysis. A systematic investigation of the detection signal for various spacer molecules between the oligonucleotide probe and the gold surface reveals that he signal/background ratio improves as the length of the spacer increases with a 12- to 18-atom spacer element being optimal. The optimal spacer molecule allows a detection limit between 30 and 100 femtomol DNA with a macroscopic gold disc electrode of 1 mm radius. The dependence of the detection signal on the concentration of a 23nt target follows a binding curve with an approximate 1:1 stoichiometry and a dissociation constant of $K_D = 13 \pm 4$ nM at 295 K.

Introduction

A possible application for a label free electrical detection system is the Avian Influenza Virus (AIV), which circulates in wild and domestic bird populations and is currently considered as a potential source for a new Influenza pandemic. Influenza viruses belong to the family of Orthomyxoviridae and are classified into types A, B and C based on the antigenic properties of their nucleoprotein and matrix protein. Further subtyping of influenza

A viruses is based on the antigenic properties of the outer surface glycoproteins haemagglutinin (HA) and neuraminidase (NA) given as HxNy subtype. So far 16 HA (H1-H16) and 9 NA (N1-N9) subtypes have been identified, although all combinations of HxNy do not occur naturally. The influenza A virus is a lipid enveloped virus, which contains a negative sense RNA genome of a total length of approximately 13.6 kBases organised into eight separate RNA molecules (segments), which code for 11 proteins [1; 2].

Influenza viruses cause a respiratory disease in humans resulting in an average death toll of 36000 people each year in the United states alone [3]. Apart from annually recurring epidemics, Influenza A viruses, which infect avian and mammalian species have been responsible for devastating pandemics killing at least 40 million people in 1918/1919 (Spanish Flu, H1N1) [4] and less serious pandemics in 1957 (Asian Influenza, H2N2), 1968 (Hong Kong Influenza, H3N2) and 1977 (Russian Influenza, H1N1) [5]. Influenza pandemics seem to occur when a highly pathogenic avian type virus acquires the capability of efficient human to human transmission (reviewed in [6]). A current threat is an avian H5N1 virus, which emerged in May 1997 [7; 8] and has caused almost 90 human deaths with the most devastating outbreak in Asia in 2003/2004 [9]. The possibility has been discussed that avian H5N1 could give rise to a new pandemic, should it acquire the capacity of efficient human to human transmission [6]. This threat makes it a necessity to detect highly pathogenic avian influenza viruses as early as possible, ideally in the field using portable detection devices.

Classical methods of influenza virus diagnostics are based on virus isolation, culture and subtype identification by immunoassays possibly followed by an in vivo experiment in order to determine pathogenicity [10]. This process requires 3 to 7 days of virus culture in addition to transporting the sample to the laboratory. Faster molecular methods based on reverse transcription polymerase chain reaction (RT-PCR) are available [11; 12; 13; 14], which require the decision prior to testing for which subtype the test should be carried out. Recently,

a cascade of real-time RT-PCRs (rRT-PCR) has been described that determines in the first step the presence of influenza A sequences by rRT-PCR specific to the matrix gene, and if positive is followed by rRT-PCRs targeted at H7- and H5-specific sequences. Where the H5 sequence is found an rRT-PCR assay has been developed that specifically detects highly pathogenic H5N1 strains of the Quinghai linage that are representative of strains currently occurring in Europe [15]. Portable RT-PCR systems have been described and applied for example to the swine fever virus [16]. As an alternative, DNA microarrays have been proposed for complete subtyping and additional virus information [17; 18; 19; 20; 21]. Current microarray technology is based on amplification of virus RNA into cDNA by RT-PCR, with the incorporation of fluorescently-labelled bases into the cDNA in order to allow subsequent optical detection of hybridisation between labelled virus-derived target cDNA and on-chip immobilised probe oligonucleotides. Microarray technology is in principle more powerful than currently established techniques allowing complete subtype identification and even short sequence information to be obtained from limited quantities of sample [20]. However, current microarray platforms require RT-PCR amplification and fluorescence imaging of the array, which is only possible with dedicated equipment available in a diagnostic lab, precluding field –testing of domestic bird populations on a global scale.

One solution to this problem would be the development of inexpensive, portable devices incorporating a disposable sensing element. In this report we investigate the application of Faradaic electrochemical impedance spectroscopy (EIS) in order to detect avian influenza gene sequences without the need for target labelling or fluorescence imaging equipment. EIS (reviewed in [22; 23]) provides detailed information about the electrical resistance and capacitance of conducting and non-conducting surfaces immersed in an electrolyte. Changes to the surface can be measured as changes in the electrical parameters of the system. The hybridisation of target DNA to surface attached probe DNA leads to an increase in negative

charge on the surface, which prevents the access of negatively charged probe molecules, e.g. the redox couple $[Fe(CN)_6]^{3-}/[Fe(CN)_6]^{4-}$, to the surface due to electrostatic repulsion [24; 25]. Label free impedance detection of oligonucleotide hybridisation has been described e.g. for interdigitated ultramicroelectrodes [26] or for electrical microarrays using a coulostatic pulse technique [27].

The electrochemical impedance spectrum is obtained by measuring the electrical current response in dependence of the frequency of the applied voltage. In any electrical circuit comprising capacitive and resistive elements exposed to an alternating voltage there will be a phase difference between the electrical current and the applied voltage in dependence of the frequency f. This is conveniently represented by the complex impedance Z

$$Z(f) = Z' + j\,Z'', \qquad (1)$$

where j represents the imaginary unit $j = \sqrt{(-1)}$. The data are typically displayed in a plot of $-Z''$ against Z', called the Nyquist plot (fig. 1A). While this does not show the frequency dependence of the impedance directly, data points obtained at the highest frequencies are in the leftmost part of the plot and those obtained at the lowest frequencies are at the right hand side of the plot. The curve can be fitted to a model of an electrical equivalent circuit (fig. 1B) allowing interpretation of the surface properties of the device in terms of electrical circuit elements. One drawback of this approach is that interpretation of the data requires an iterative curve-fitting algorithm and user input, which is not readily accomodated in a mobile device.

For oligonucleotide probes attached via –SH groups to the surface of gold electrodes our results show significant changes in the EIS curves upon hybridisation with target sequences. The quality and magnitude of the detection signal depends on the nature and length of various spacers between the –SH group and the oligonucleotide. We devise a novel procedure for data analysis, which does not require curve-fitting or human intervention and therefore is

particularly suited for a portable field device. The detection limit for a 23 nucleotides (nt) long DNA sequence derived from a highly pathogenic H5N1 AIV strainis between 30 and 100 fmol.

Materials and Methods

Selection of oligonucleotides

Virus gene sequences were obtained from the NCBI influenza virus resource [28] and further processed with BioEdit software (version 7.0.5.3) [29; 30]. Multiple sequence alignment was performed with MUSCLE (version 3.6) [31; 32].

The majority of sulfhydryl modified oligonucleotide probes were obtained from Eurogentec (Seraing, Belgium) and unmodified target oligonucleotides from MWG Biotech (London, UK). The probes MA20/S18/C6 and HA23/S18/C6 were obtained from the Protein & Nucleic Acid Chemistry Facility (Department of Biochemistry, University of Cambridge, UK).

Preparation of electrodes

Gold disk working electrodes with a radius of 1 mm were obtained from CH Instruments (Austin TX, USA) and polished with aluminium oxide particles of 5 μm diameter (Buehler, Lake Bluff, Il, USA) followed by sonication in water using a conventional sonicator (150 W power output), then with 1 μm particles followed by sonication in water/detergent, then with 0.3 μm particles followed by sonication in water and finally polished with a soft pad (Buehler) followed by sonication in water. They were further cleaned electrochemically in 0.5 M H_2SO_4 by cyclic voltammetry varying the potential between -0.05 V and +1.1 V

against a Hg/Hg_2SO_4 reference electrode for 60 cycles using a three-electrode cell with the gold electrode as the working electrode and a platinum wire as the counter electrode. The electrodes were connected to an Autolab PGSTAT302 potentiostat (Eco Chemie, Utrecht, Netherlands). For SAM formation and oligonucleotide immobilisation the electrode was dried briefly in a stream of nitrogen and then incubated for 20 hours in a hydration chamber at 22°C with 20 µL of a buffer solution of 100 mM Tris-hydroxymethyl-aminoethane, pH 8.5, 100 mM NaCl, 5 mM $MgCl_2$ containing 10 µM mercaptohexanol and 1 µM sulfhydryl modified DNA oligonucleotide (fig. 2). The electrode was then rinsed for 2 min in rinse buffer (100 mM phosphate buffer pH 7.2, 100 mM NaCl, 5 mM $MgCl_2$) before measurement of the impedance spectrum. The impedance spectrum was measured with the Autolab PGSTAT302 in a solution of 5 mM $K_4[Fe(CN)_6]$, 5 mM $K_3[Fe(CN)_6]$ and 100 mM phosphate buffer at pH 7.2 using the three-electrode cell described above, while the reference electrode was connected via a salt bridge filled with 100 mM phosphate buffer. The impedance spectrum was measured in the frequency range from 100 kHz to 100 mHz, while a potential of -0.195 V versus the reference electrode was applied, which corresponds to the formal potential of the redox couple.

For backfilling the electrode was treated for 2 hours in 20 µL of a solution of 1 mM mercaptohexanol, 100 mM phosphate buffer pH 7.2, 100 mM NaCl, 5 mM $MgCl_2$ in a hydration chamber at 22°C followed by measurement of the impedance spectrum as described above.

Sensing of target oligonucleotides

The prepared gold electrode surface was covered with 20 µL of a solution of target oligonucleotide of variable concentration (see results section) in 100 mM phosphate buffer pH 7.2, 100 mM NaCl, 5 mM $MgCl_2$ for 2 hours at 22°C in a hydration chamber. The minimum amount of solution necessary to cover the surface was estimated to be 10 µL,

although 20 μL has been used always. Then the electrode was rinsed for two minutes in the same buffer without oligonucleotide and the impedance spectrum was measured as described above.

Data analysis

Curve fitting of the impedance spectra to an electrical equivalent circuit was performed with the Frequency Response Analyser software supplied with the Autolab potentiostat (Eco Chemie, Utrecht, Netherlands). Only impedances obtained at frequencies between 0.1 Hz and 10 kHz were used for curve fitting in order to avoid contributions from the electrochemical cell and connecting wires at high frequencies and the influence of DC-conductivity of the electrolyte solution at low frequencies. The normalised hybridisation signal was obtained by converting the Autolab data files into space delimited ASCII files and subsequent calculation of the signal using a computer program developed in our lab.

The concentration dependence of the signal S was fitted with a binding curve for the reaction between probe P and target R, i.e. $P + r\,T \rightarrow PT_r$

$$S = \frac{S_{\max}[\mathrm{cHA23}]^r}{K_D^r + [\mathrm{cHA23}]^r} \qquad (4)$$

where S_{max} denotes the maximum signal for $[cHA23] \rightarrow \infty$, r is the stoichiometric factor and K_D is the dissociation constant of PT_r.

Results

Selection of probe- and target-oligonucleotide sequences

Highly pathogenic avian influenza viruses have been found to possess several basic amino acids inserted into the posttranslational cleavage site of the HA precursor protein [33]. Therefore, we selected an amino acid sequence typical of highly pathogenic viruses,

PQRERRRKKR, which occurs in 498 of 8051 available HA sequences from the NCBI influenza Resource [28]. In particular this sequence occurs in an avian H5N1 virus isolated from a human in 1997, A/Hong Kong/503/97, GenBank accession number AF102679. The reverse translation of this sequence into DNA yields the 30 nucleotides (nt) sequence 5′-cctcaaagagagaagaagaaaaaagagagga-3′, which is found in 438 H5N1 sequences from virus isolates between 1997 and 2006. In order to detect a larger number of virus isolates we truncated the sequence to the 23nt fragment 5′-cctcaaagagagaagaagaaa-3′ (designated here HA23), which occurs in 586 virus isolates [28], potentially enabling a future biosensor to detect a larger number of highly pathogenic viruses. This sequence, however, has 101 significant matches in the 'whole genome shotgun sequence' database of the NCBI, thus positive detection events due to other genes in the environment cannot be excluded with a single probe. In order to detect influenza A viruses specifically, we selected a 20 nt sequence from an invariant portion of the matrix gene (MA), gtgagcgaggactgcagcgt (MA20), which also has a number of matches with non-influenza genes but most importantly no double matches with the HA23 sequence. In order to model negative strand influenza virus RNA, we designed two short target oligonucleotides cHA23 and cMA20, which are the reverse complement to the probe sequences HA23 and MA20 (table 1). A more realistic scenario encountered in environmental samples will be longer target RNA fragments, likely to be generated by sample degradation. Thus we designed a 120nt oligonucleotide derived from the haemagglutinin gene, which contains 23 nucleotides reverse complementary to HA23 (cHA120, table 1).

Electrochemical characterisation of the biosensing event

The Faradaic impedance spectra of gold disc electrodes with a radius of 1 mm were obtained in the presence of the negatively charged redox couple $[Fe(CN)_6]^{3-}/[Fe(CN)_6]^{4-}$ at various

stages of the electrode preparation process (fig. 1A). The spectra of modified electrodes show two components, a semi-circle followed by a linear increasing portion. The unmodified gold electrode showed a barely visible semicircle at higher frequencies followed by a linear part which covers most of the spectrum. Overnight treatment of the gold electrode with an aqueous buffer containing a mixture of 10 μM mercaptohexanol and 1 μM of an –SH labelled oligonucleotide (HA23/S12/C6, see fig. 2) at pH 8.5 resulted in a clear change of the curve, with a significant semicircle portion, followed by a linear increasing portion. The curve did not change upon rinsing with buffer, which indicates the formation of a mixed oligonucleotide/mercaptohexanol self-assembled monolayer (SAM) on the gold surface. Subsequent treatment of the gold surface with 1 mM mercaptohexanol for two hours ("backfilling") resulted in a reduction of the semicircle diameter, but overall a relatively small change was observed.

For the biosensing event the electrode was incubated for two hours with a reverse complementary target cHA23 (see table 1). The Nyquist plot (fig. 1A) showed a significant change of the curve caused by an increased diameter of the semicircle compared to the plot obtained before treatment with the target oligonucleotide.

The data have been modelled with a Randles and Ershler electronic equivalent circuit [34; 35] shown in figure 1B. R_1 represents the ohmic resistance of the electrolyte solution and the Warburg impedance W results from the diffusion of ions from the bulk electrolyte to the gold surface. The constant phase element (CPE) represents the combined capacitance of the SAM and the electrochemical cell and R_{ct} is the resistance for charge transfer from solution to the electrode caused by the electrochemical double layer and the negatively charged oligonucleotides repelling the negatively charged redox probe molecule. The use of a CPE instead of a capacitor resulted in better fits, likely because the CPE represents the non-homogeneity of the electrode surface. The extent of deviation from an ideal capacitor can be

modelled by a parameter n ($0 <= n <= 1$), where n = 1 for an ideal capacitor [36]. The parameters of the fit (table 2) show that the biosensing of DNA hybridisation was characterised by a marked increase in the electron transfer resistance from R_{ct} = 2.78 kΩ to 4.85 kΩ i.e. a 75% increase, while the parameters of the CPE changed slightly. The calculated capacitance decreases slightly. The pattern of relative change in the parameters for each electrode upon hybridisation with a complementary DNA oligonucleotide was similar for all experiments with the same spacer molecule, but the electrochemical characteristics of the SAM before hybridisation shows large variations in the electron transfer resistance from R_{ct} = 1.0 kΩ to R_{ct} = 7.0 kΩ.

Treatment of the SAM with a non-complementary DNA sequence resulted in an insignificant shift of the curve in the Nyquist plot (fig. 4A). However, the change in electrical signal in detail upon hybridisation or upon treatment with a non-complementary sequence depended on the nature and length of the spacer molecule (fig. 2) between the oligonucleotide probe and the gold surface, which is investigated below.

In a similar experiment we demonstrated the detection of a gene sequence from the influenza matrix protein cMA20 using the 20mer oligonucleotide MA20/S18/C6 as a probe attached to the gold surface. The charge transfer resistance changed from R_{ct} = 6.59 kΩ to R_{ct} = 9.89 kΩ, i.e. by 50% upon treatment with cMA20 (fig. 3A), while the charge transfer resistance increased only from R_{ct} = 5.38 kΩ to R_{ct} = 5.98 kΩ for a different electrode treated with cHA23 (fig. 3B).

Automated data analysis

The curve-fitting of the impedance spectra illustrated above requires a complex fitting algorithm and human intervention for the setting of starting values. For a biosensor which can be used in the field, an automatic procedure for obtaining the signal using an algorithm with a minimal amount of calculations is desirable. We achieve this by using the modulus of the

impedance $|Z|_f$ in dependence of frequency f of the alternating voltage. From the complex impedance defined in eq. (1) the modulus is obtained as

$$|Z|_f = \sqrt{Z'^2 + Z''^2} \qquad (2)$$

The hybridisation signal is calculated from the sum over N data points of the absolute differences in |Z| before and after hybridisation:

$$\text{Signal} = \frac{\sum_f Abs(|Z|_f^{after} - |Z|_f^{before})}{\max Z''_{before} \cdot N} \qquad (3)$$

In order to account for all available data, the sum is taken over all frequencies of the measurement. In order to account for differences between individual electrodes the signal is and normalised by the number of data points N and the maximum of the phase shift in the semicircle portion of the impedance spectrum before hybridisation. The procedure is illustrated in figure 4 for two experiments using a non-complementary target sequence (fig. 4A) and a complementary target sequence (fig. 4B). The star denotes the normalisation factor $\max Z''_{before}$, which is related to the electron transfer resistance of the electronic equivalent circuit shown in fig. 1B, $\max Z'' = \frac{1}{2} R_{ct}$ [22]. The formation of the cumulative sum in dependence of frequency is illustrated in fig. 4C for the detection event (upper curve) and for a non-detection event (lower curve) giving a signal ≈80 kΩ for the detection event and ≈4 kΩ for a non-detection event. The normalised unitless signal is 0.9296 and 0.0447 respectively. For the following experiments the normalised signal is reported solely.

The influence of spacer length

The oligonucleotide probe HA23 was attached at the end of each spacer, which was attached via a 6-carbon chain to a sulfhydryl group as shown in fig. 2. The 6-carbon chain is of the same length than mercaptohexanol, which was used for backfilling in order to block non-specific interactions with the gold electrode. For each spacer the biosensor signal was

obtained by incubation of the probe with either fully complementary target (cHA23, table 1) or non-complementary target cMA20. The average biosensor signal from at least 3 experiments for all spacer molecules varies from 0.20 ± 0.11 to 0.71 ± 0.30 for treatment with complementary target and from 0.021 ± 0.032 to 0.12 ± 0.08 for non-complementary target. The average results shown in fig. 5 reveal a clear change of the signal for complementary target, whose magnitude varies with spacer length. In contrast, the signal for non-complementary target is similar for all spacer lengths investigated. Some spacer molecules are clearly unsuitable for biosensing of oligonucleotide hybridisation, e.g. HA23/S3/C6. In contrast, HA23/S12/C6 gives the highest signal followed by HA23/S18/C6which also gives the highest ratio of specific/non-specific signal (Figure 5). However, only in a limited number of cases were we able to obtain stable SAMs with the HA23/S18/C6 probe, possibly indicating problems in synthesis of the longer spacer element, and thus we chose HA23/S12/C6 for further experiments.

Dependence on target concentration

Figure 6A shows the dependence of the biosensing signal for the HA23/S12/C6 probe in dependence of the complementary target concentration cHA23. The dashed line indicates the level of signal after treatment with non-complementary sequence at a concentration of 1000 nM. The detection limit is between 3 nM and 10 nM cHA23, when the signal exceeds the signal for non-complementary sequence. The signal increases steeply between 3 nM and 30 nM, while between 30 nM and 1000 nM there is no clear increase in signal, but the fluctuations are within the error margin. The curve has been fitted with a classical binding curve (see methods) yielding a maximum signal of $S_{max} = 0.70 \pm 0.01$, while the phenomenological dissociation constant is $K_D = 13.4 \pm 4$ nM and the exponent is $r = 1.31 \pm 0.08$, which is close to a value expected for a 1:1 complex. The dissociation constant

compares well with K_D = 6.8 nM obtained from the on/off rate constant for hybridisation of a 19 nt oligonucleotide target to a 19 nt surface attached probe [37].

The detection limit is between 3 and 10 nM target molecules in 20µL solution, the volume used to cover the gold electrode surface, corresponding to 60 to 200 fmol of target molecules.

Detection of long target sequences

In order to test the ability of the biosensor to detect larger gene fragments as they may occur in samples taken from infected birds, we investigated the detection of a 120mer target sequence containing 23 nucleotides reverse complementary to the probe HA23/S12/C6 (fig. 6B). The average signal obtained with 1000 nM cHA120 target is 0.56, a value lower than obtained with the cHA23. The detection limit is in the region of 10 nM target concentration corresponding to 200 fmol of target molecules.

Discussion

The appearance of a prominent semi-circle in the electrochemical impedance spectrum (EIS) upon treatment of the gold electrode with mercaptohexanol/HS-oligonucleotide (1:10) is indicative of the formation of a self-assembled monolayer (SAM). It has been shown that native DNA has a high affinity for gold and may adsorb non-specifically to the gold surface [38]. In order to reduce non-specific adsorption we treated the gold electrode with a buffer containing a high concentration of mercaptohexanol (backfilling) as well as using a 10:1 molar ratio of mercaptohexanol to probe DNA. Other molar ratios did not lead to a successful SAM formation (see additional information). The reduction of the charge-transfer resistance upon backfilling (fig. 1A) indicates that backfilling is effective in reducing non-specific adsorption in addition to the use of mercaptohexanol in the initial step of electrode

preparation. For samples taken in the field, a nucleic acid extraction step may be included, that reduces the amount of protein and cellular material that may bind non-specifically to gold electrodes. Alternatively an ethyleneglycol-capped thiol could be used, which has been reported to have very low protein and cellular adsorptivity [39]. The large variation in charge transfer resistance between various electrodes is caused by the manual steps of polishing (see methods) and variations in the gold surface between electrodes, which lead to a variation in surface properties at the molecular level thus influencing SAM formation. However, for the detection of hybridisation only the change of the curve or the magnitude of the derived detection signal is relevant. The devised biosensing system unequivocally demonstrates the detection of DNA sequences complementary to avian influenza gene probes. In particular, we propose to use a 23nt probe characteristic of high virulence pathogenic strains (HA23) and a 20nt probe characteristic of all influenza viruses (MA20), which when taken together, would clearly identify a highly pathogenic influenza virus, despite the presence of genes from other organisms that might be present in an environmental sample. For a mobile device with limited computational power used by an unskilled operator, the automated data analysis resulting in a single numerical output (the signal) will be useful. Additionally the automated data analysis does not require a model of an electrical equivalent circuit, which might not be valid for all types of electrodes, even if they are manufactured to a high standard. The model free data analysis described here yields a signal whenever there is a change in the EIS.

In order to optimise the detection system the influence of various spacer molecules shown in fig. 2 on the detection signal was investigated. The results in fig. 5 show that some spacer molecules are less suitable or completely unsuitable (e.g. HA23/S3/C6), as the error margin of the detection signal with complementary target overlaps with the error margin of the signal in presence of non-complementary target. HA23/S12/C6 gives the highest detection signal

followed by HA23/S18/C6. In general longer spacer molecules give a higher hybridisation signal. The reason for this may be the increased mobility of the probe, which makes it more readily available for hybridisation and once hybridised the resulting highly charged duplex DNA is able to cover a wider area of the surface leading to increased electrostatic repulsion of the negatively charged redox marker. Interestingly the error margins between detection of complementary sequence and non-complementary sequence are further apart for HA23/S18/C6 than for HA23/S12/C6, which would render this spacer as equally suitable for a detection device. It would be interesting to investigate, whether longer spacers lead to further improvements, although the chemistry is currently not available to us.

The concentration dependence of the detection signal shows a detection limit below 200 femtomols of target molecule, with a macroscopic gold electrode using a buffer volume of 20 μL in order to cover the gold surface. For the current experimental setup the volume could be reduced to 10 μL, thus reducing the detection limit of the number of target molecules even further. As the detection limit depends on the area of the electrodes we would expect a lower detection limit for electrodes used in a microarray device. Additionally much lower sample volumes are required to cover cover a microelectrode. Current DNA-microarrays for optical detection use volumes in the picolitre range, e.g. for 100 pL, the theoretical detection limit would be reduced to 1 attomol, i.e. approximately $6 \cdot 10^5$ molecules, which compares well with typical viral loads of 10^5 to 10^6 copy numbers found in oropharyngeal samples [40]. Additionally the current hybridisation time of 2 h would be reduced. However, the EIS for smaller electrodes may be limited by noise effects and contributions from the edges of electrodes, which become more prominent if the electrode surface is reduced, thus amplification by RT-PCR may be necessary.

The detection limit of the currently most sensitive rRT-PCR method has been reported to be between 5 and 50 RNA copies per reaction [12], which is substantially lower than the detection limit reported in this work. However, compared to a multiplex rRT-PCR system, a microarray EIS system would have the advantage of a smaller size, provide more information, provide faster read-out times and potentially possess a disposable sensor element, which reduces the risk of spreading the virus.

The concentration dependence of the signal follows a classical binding curve, while the exponent in region of $r \approx 1$ indicates a 1:1 complex formation between probe and target as would be expected for the formation of a DNA duplex from single stranded oligonucleotides. However, the binding curve is a phenomenological description of the concentration dependence of the signal. The actual binding may be a stepwise process, in which shorter stretches of base-pairs are formed in succession explaining the non-integer value of $r = 1.3$ obtained.

The detection signal for a short (23nt) oligonucleotide is higher than for a longer oligonucleotide (120nt) containing the 23nt sequence within it. This might be caused by the space constraints on the sensor surface. Once the DNA duplex is formed, non-hybridised stretches about 50 nucleotides long of DNA remain on each side, which may prevent hybridisation at neighbouring probe molecules. Additionally the hybridisation efficiency of such a long target may be lower than for a short completely matched target taking into account the 50 nucleotide overhang on the 5' end close the gold surface. This would also explain why the detection limit of at least 10 nM is higher for the 120mer target than for the 23mer target. These observations have strong implications for the design of DNA microarrays. Although longer oligonucleotide probes may confer greater specificity, shorter oligonucleotides may give rise to greater sensitivity. Because the microarray format allows

specificity to be determined from combinations of oligonucleotide probes as discussed here, shorter oligonucleotides may prove to be the probes of choice. There are also implications for sample preparation: techniques designed to maintain sample (RNA, DNA) integrity may be detrimental to signal generation, if short ligands are preferred.

Conclusions

An optimized procedure for SAM formation on gold electrodes has been used, which employs a 1: 10 DNA/mercaptohexanol ratio, followed by back filling with mercaptohexanol alone (see Additional Material). A systematic investigation of the spacer molecule between the oligonucleotide and the gold surface identifies HA23/S12/C6 or HA23/S18/C6 (see fig. 2) as the optimal length.

With these probes we have shown the detection of a 23 nt DNA sequence as a model for an avian influenza gene sequence to a detection limit of 200 femtomol and below. These results form the basis for the development of a portable gold-microelectrode array device using the impedimetric detection principle, which is able to analyse environmental samples for the presence of highly pathogenic avian influenza virus strains that may present a risk to humans.

Acknowledgements

The authors wish to thank Dr Michael Helmus of Advance Nanotech for pointing out the importance of label free detection in the field of Avian Influenza and for helpful discussions. Special thanks goes to Simon Keighley for optimising the procedure for the preparation of electrodes, assisting with measurement and helpful suggestions for the current manuscript. This work has been supported by Advance Nanotech, in partnership with the Centre for Advanced Photonics and Electronics (CAPE) of the Department of Engineering, University

of Cambridge, under the BiMAT project. PE

acknowledges support from the BBSRC under contract BB/D523094/1. Work in the PKF lab was supported by a Grant in Aid from the Medical Research Council to the Cancer Cell Unit. AK and PKF acknowlege the generous support of the Hutchison/MRC Research Centre and Cancer Cell Unit, where some of the work has been performed.

References

[1] J.C. Obenauer, J. Denson, P.K. Mehta, X. Su, S. Mukatira, D.B. Finkelstein, X. Xu, J. Wang, J. Ma, Y. Fan, K.M. Rakestraw, R.G. Webster, E. Hoffmann, S. Krauss, J. Zheng, Z. Zhang, and C.W. Naeve, Large-scale sequence analysis of avian influenza isolates. Science 311 (2006) 1576-80.

[2] D. Kobasa, and Y. Kawaoka, Emerging influenza viruses: past and present. Curr Mol Med 5 (2005) 791-803.

[3] J. Li, S. Chen, and D.H. Evans, Typing and subtyping influenza virus using DNA microarrays and multiplex reverse transcriptase PCR. J Clin Microbiol 39 (2001) 696-704.

[4] N.P. Johnson, and J. Mueller, Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. Bull Hist Med 76 (2002) 105-15.

[5] N.J. Cox, and K. Subbarao, Global epidemiology of influenza: past and present. Annu Rev Med 51 (2000) 407-21.

[6] T. Horimoto, and Y. Kawaoka, Influenza: lessons from past pandemics, warnings from current incidents. Nat Rev Microbiol 3 (2005) 591-600.

[7] E.C. Claas, A.D. Osterhaus, R. van Beek, J.C. De Jong, G.F. Rimmelzwaan, D.A. Senne, S. Krauss, K.F. Shortridge, and R.G. Webster, Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. Lancet 351 (1998) 472-7.

[8] K. Subbarao, A. Klimov, J. Katz, H. Regnery, W. Lim, H. Hall, M. Perdue, D. Swayne, C. Bender, J. Huang, M. Hemphill, T. Rowe, M. Shaw, X. Xu, K. Fukuda, and N. Cox, Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. Science 279 (1998) 393-6.

[9] W.H. Organisation, Avian Influenza A (H5N1), Weekly Epidemiology Rev., 2004, pp. 65-70.

[10] G.A. Storch, Essentials of diagnostic virology, Churchill Livingstone, New York, 2000.

[11] S. Payungporn, S. Chutinimitkul, A. Chaisingh, S. Damrongwantanapokin, C. Buranathai, A. Amonsin, A. Theamboonlers, and Y. Poovorawan, Single step multiplex real-time RT-PCR for H5N1 influenza A virus detection. J Virol Methods 131 (2006) 143-7.

[12] L. Di Trani, B. Bedini, I. Donatelli, L. Campitelli, B. Chiappini, M.A. De Marco, M. Delogu, C. Buonavoglia, and G. Vaccari, A sensitive one-step real-time PCR for detection of avian influenza viruses using a MGB probe and an internal positive control. BMC Infect Dis 6 (2006) 87.

[13] E. Spackman, D.A. Senne, T.J. Myers, L.L. Bulaga, L.P. Garber, M.L. Perdue, K. Lohman, L.T. Daum, and D.L. Suarez, Development of a real-time reverse

transcriptase PCR assay for type A influenza virus and the avian H5 and H7 hemagglutinin subtypes. J Clin Microbiol 40 (2002) 3256-60.
[14] S. Payungporn, P. Phakdeewirot, S. Chutinimitkul, A. Theamboonlers, J. Keawcharoen, K. Oraveerakul, A. Amonsin, and Y. Poovorawan, Single-step multiplex reverse transcription-polymerase chain reaction (RT-PCR) for influenza A virus subtype H5N1 detection. Viral Immunol 17 (2004) 588-93.
[15] B. Hoffmann, T. Harder, E. Starick, K. Depner, O. Werner, and M. Beer, Rapid and Highly Sensitive Pathotyping of Avian Influenza A H5N1 Virus by Using Real-Time Reverse Transcription-PCR. J Clin Microbiol 45 (2007) 600-3.
[16] G.R. Risatti, J.D. Callahan, W.M. Nelson, and M.V. Borca, Rapid detection of classical swine fever virus by a portable real-time reverse transcriptase PCR assay. J Clin Microbiol 41 (2003) 500-5.
[17] E.D. Dawson, C.L. Moore, J.A. Smagala, D.M. Dankbar, M. Mehlmann, M.B. Townsend, C.B. Smith, N.J. Cox, R.D. Kuchta, and K.L. Rowlen, MChip: a tool for influenza surveillance. Anal Chem 78 (2006) 7610-5.
[18] M.B. Townsend, E.D. Dawson, M. Mehlmann, J.A. Smagala, D.M. Dankbar, C.L. Moore, C.B. Smith, N.J. Cox, R.D. Kuchta, and K.L. Rowlen, Experimental evaluation of the FluChip diagnostic microarray for influenza virus surveillance. J Clin Microbiol 44 (2006) 2863-71.
[19] A.L. Ghindilis, M.W. Smith, K.R. Schwarzkopf, K.M. Roth, K. Peyvan, S.B. Munro, M.J. Lodes, A.G. Stover, K. Bernards, K. Dill, and A. McShea, CombiMatrix oligonucleotide arrays: Genotyping and gene expression assays employing electrochemical detection. Biosens Bioelectron (2006).
[20] R.H. Liu, M.J. Lodes, T. Nguyen, T. Siuda, M. Slota, H.S. Fuji, and A. McShea, Validation of a fully integrated microfluidic array device for influenza a subtype identification and sequencing. Anal Chem 78 (2006) 4184-93.
[21] M.J. Lodes, D. Suciu, M. Elliott, A.G. Stover, M. Ross, M. Caraballo, K. Dix, J. Crye, R.J. Webby, W.J. Lyon, D.L. Danley, and A. McShea, Use of semiconductor-based oligonucleotide microarrays for influenza a virus subtype identification and sequencing. J Clin Microbiol 44 (2006) 1209-18.
[22] E. Katz, and I. Willner, Probing biomolecular interactions at conductive and semiconductive surfaces by impedance spectroscopy: routes to impedimetric immunsensors, DNA-sensors and enzyme biosensors. Electroanalysis 15 (2003) 913-947.
[23] H.G.L. Coster, T.C. Chilcott, and A.C.F. Coster, Impedance spectroscopy of interfaces, membranes and ultrastructures. Bioelectrochemistry and Bioenergetics 40 (1996) 79-98.
[24] H. Aoki, P. Bühlmann, and Y. Umezawa, Electrochemical detection of a one-base mismatch in an oligonucleotide using ion-channel sensors with self-assembled PNA monolayers. Electroanalysis 12 (2000) 1272-1276.
[25] S. Nagase, M. Kataoka, R. Nanganawa, R. Komatsu, K. Odashima, and Y. Umezawa, Voltammetric anion responsive sensors based on modulation of ion permeability through Langmur-Blodgett films containing synthetic anion receptors. Analytical Chemistry 62 (1990) 1252-1259.
[26] V. Dharuman, T. Grunwald, E. Nebling, J. Albers, L. Blohm, and R. Hintsche, Label-free impedance detection of oligonucleotide hybridisation on interdigitated ultramicroelectrodes using electrochemical redox probes. Biosens Bioelectron 21 (2005) 645-54.

[27] V. Dharuman, E. Nebling, T. Grunwald, J. Albers, L. Blohm, B. Elsholz, R. Worl, and R. Hintsche, DNA hybridization detection on electrical microarrays using coulostatic pulse technique. Biosens Bioelectron 22 (2006) 744-51.
[28] Influenza Virus Resource, National Center for Biotechnology Information, 2004.
[29] T.A. Hall, Bioedit, 2005.
[30] T.A. Hall, BioEdit: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41 (1999) 95-98.
[31] R.C. Edgar, MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5 (2004) 113.
[32] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32 (2004) 1792-7.
[33] A. Stieneke-Grober, M. Vey, H. Angliker, E. Shaw, G. Thomas, C. Roberts, H.D. Klenk, and W. Garten, Influenza virus hemagglutinin with multibasic cleavage site is activated by furin, a subtilisin-like endoprotease. Embo J 11 (1992) 2407-14.
[34] J.E.B. Randles, Discuss Faraday Soc 1 (1947) 11.
[35] B.V. Ershler, Discuss Faraday Soc 1 (1947) 269.
[36] A.J. Bard, and L.R. Faulkner, Electrochemical Methods: Fundamentals and Applications, Wiley, New York, 1980.
[37] D. Li, X. Zou, Q. Shen, and S. Dong, Kinetic study of DNA/DNA hybridisation with electrochemical impedance spectroscopy. Electrochemistry Communications 9 (2007) 191-196.
[38] A.B. Steel, R.L. Levicky, T.M. Herne, and M.J. Tarlov, Immobilization of nucleic acids at solid surfaces: effect of oligonucleotide length on layer assembly. Biophys J 79 (2000) 975-81.
[39] E. Ostuni, R.G. Chapman, M.N. Liang, G. Meluleni, G. Pier, D.E. Ingeber, and G.M. Whitesides, Self-Assembled Monolayers That Resist the Adsorption of Proteins and the Adhesion of Bacterial and Mammalian Cells. Langmuir 17 (2001) 6336-6343.
[40] H. Toro, D.C. Tang, D.L. Suarez, M.J. Sylte, J. Pfeiffer, and K.R. Van Kampen, Protective avian influenza in ovo vaccination with non-replicating human adenovirus vector. Vaccine 25 (2007) 2886-91.
[41] C.S. Hsu, and F. Mansfield, Concerning the conversion of the constant phase element parameter Y0 into a capacitance. Corrosion 57 (2001) 747.

Figures

Figure 1 - Nyquist plot of typical impedance spectra

A) Nyquist plots after various stages of electrode preparation, the bare gold surface (♦), after the preparation of the SAM (□), after further treatment with mercaptohexanol (backfilling) (○) and after hybridisation with cHA23 (Δ). The solid lines show the curve-fit of the electrical equivalent circuit in B) to the data.

B) Model electrical equivalent circuit to describe the data shown in A). R_1 represents the electrolyte resistance, R_{ct} the charge-transfer resistance, CPE the constant phase element and W the Warburg impedance.

Figure 2 - Probe molecules

Chemical structure of the probe molecules and the backfilling agent mercaptohexanol used in this study with the abbreviations (left had side) used in the text. NNN… denotes the single stranded oligonucleotide sequence. The sulfhydryl group is always at the end of a 6 carbon chain, at the end of which is attached the variable spacer. The spacer is formed either by pure carbon chains (C_3 or C_{12}) or mixed carbon/oxygen chains of 9 or 18 atoms in total.

Figure 3 - Detection of cMA20

A) EIS of a gold electrode with MA20/S18/C6 probes before (□) and after (○) treatment with reverse complementary target cMA20. The signal obtained from the model-free data analysis is given.

B) EIS of a gold electrode with MA20/S18/C6 probes before (□) and after (○) treatment with non-complementary target cHA23.

Figure 4 - Illustration of model-free data analysis

A) EIS before (○) and after (●) treatment of the sensor with a non-complementary target. B) EIS before (○) and after (●) treatment with a complementary target. C) Frequency dependence of the cumulative sum of the differences in the modulus |Z| of the EIS curves for treatment with a complementary target (upper curve) and treatment with a non-complementary target (lower curve). The arrows point at the magnitude of the signal before normalisation. For normalisation the signal is divided by the number of data points and by the maximum of –Z'' in the semicircle part of the EIS highlighted with * in A) and B).

Figure 5 - Influence of spacer molecules

The dependence of the signal for the HA23 oligonucleotide attached to the gold surface with various spacer molecules treated with a complementary target and a non-complementary target. Signal averages are shown of at least three experiments. The ratio of specific/non-specific signal is given for each spacer molecule.

Figure 6 - Concentration dependence

A) Dependence of the detection signal on the target concentration of cHA23 for the probe HA23/S12/C6. Averages of at least 3 experiments are shown. The line of best fit has been obtained from the binding curve described in equation 4.

B) Demonstration of the ability to detect the 120mer target cHA120 with the probe HA23/S12/C6 at 1000 nM target concentration. The level of signal for lower target concentrations is shown in addition. Note that error bars for concentrations below 1000 nM are not displayed for this proof-of-concept experiment.

The dashed lines indicate the level of signal obtained with 1000 nM of the non-complementary target cMA20.

Tables

Table 1 - DNA oligonucleotide probes and targets

Name	Function	Sequence (5' –> 3')
HA23	Probe for pathogenic avian H5N1 haemagglutinin gene sequence	5′-cctcaaagagagagaagaagaaa-3′
MA20	Probe for influenza A matrix protein gene sequence	5′-gtgagcgaggactgcagcgt-3′
cHA23	Target, reverse complement to HA23	5′-tttcttcttctctctctttgagg-3′
cHA120	Long target sequence with 23 nucleotides reverse complementary to HA23	5′-ccatcctcccctctataaaacctgc tatagctccaaataatcctctct **ttttttcttcttctctctctttga gg**gctatttctgagcccagtcgcaa ggactaatctgtttgatttcacata-3′
cMA20	Target, reverse complement to MA20	5′-acgctgcagtcctcgctcac-3′

Table 2 - Fit parameters for curves in figure 1

Experiment	R_1/Ω (± 2%)	CPE/(μS s^n) (± 5%)	n (± 1%)	W/(mS $s^{1/2}$) (± 8%)	$R_{ct}/k\Omega$ (± 2%)	$C_{calc}/\mu F^*$
SAM preparation	150	0.8879	0.8892	0.5166	3.067	0.4254
Backfilling	156	0.5962	0.9176	0.7104	2.776	0.3355
Hybridisation	143	0.5585	0.9235	0.6716	4.85	0.3423

The meaning of the variables is explained in figure 1.
*) The capacitance has been calculated according to the relationship $C_{calc} = (CPE\ R_{ct})^{1/n}/R_{ct}$ [41].