Nanohole array plasmonic biosensors: Emerging point-of-care applications

Alisha Prasad\textsuperscript{a}, Junseo Choi\textsuperscript{a,b}, Zheng Jia\textsuperscript{a,b}, Sunggook Park\textsuperscript{a,b,*}, Manas Ranjan Gartia\textsuperscript{a,**}

\textsuperscript{a} Department of Mechanical and Industrial Engineering, Louisiana State University, Baton Rouge, LA 70803, USA
\textsuperscript{b} NIH Center for BioMolecular Multiscale Systems for Precision Medicine, Louisiana State University, Baton Rouge, LA 70803, USA

\textbf{A R T I C L E   I N F O}

Keywords: Nanohole array Category of sensors Geometrical parameters Micro/nano-fabrication Plasmonic biosensors Point-of-care biosensing

\textbf{A B S T R A C T}

Point-of-care (POC) applications have expanded hugely in recent years and is likely to continue, with an aim to deliver cheap, portable, and reliable devices to meet the demands of healthcare industry. POC devices are designed, prototyped, and assembled using numerous strategies but the key essential features that biosensing devices require are: (1) sensitivity, (2) selectivity, (3) specificity, (4) repeatability, and (5) good limit of detection. Overall the fabrication and commercialization of the nanohole array (NHA) setup to the outside world still remains a challenge. Here, we review the various methods of NHA fabrication, the design criteria, the geometrical features, the effects of surface plasmon resonance (SPR) on sensing as well as current state-of-the-art of existing NHA sensors. This review also provides easy-to-understand examples of NHA-based POC biosensing applications, its current status, challenges, and future prospects.

1. Introduction

Nanohole arrays (NHAs) are a class of nanostructured material consisting of nanoscale voids fabricated on the surface of a metallic material (O’Mahony, 2011). NHA architectures have been accepted as sub-diffraction optical systems with high aspect-ratio that involves extraordinary optical transmission (EOT) and large electric field intensities. With the discovery of EOT by (Ebbesen et al., 1998a) on metallic substrates with sub-wavelength holes, EOT has ever since been an area of major research. Several structures with EOT have been designed for plasmonics and nano-photonic. EOT occurs due to the excitation of surface plasmons (SPs) generating several scattering orders that are reliant on the shape, periodicity of nanoholes, their dielectric properties, and composition of materials (Najiminaini et al., 2012). The decoupled light from SPs radiates on the side opposite to the metal layer and spawns localized surface plasmon resonances (LSPRs) aiding to fascinating optical properties with promising applications.

As shown in Fig. 1a, in SPR sensor, the molecules are detected based on the changes in the refractive index (RI) within an evanescent field of the SPs near the sensor surface. These SPs are quickly attenuated by the metal film which is a major limitation of SPR based sensors. To overcome this, various strategies have been employed. For example, classical waveguide interferometry (CWI) utilizes electromagnetic waves to extract the information over the whole sensor surface improving the binding actions and sensitivity (Fig. 1b). Similar to CWI, Grating Coupled Interferometry (GCI) is another type of optical label free sensor, which overcomes the alignment issues of CWI, such that the reference beam is also coupled into the waveguide as shown in Fig. 1c. Few advantages of GCI over conventional SPR and CWI sensors are improved performance due to robust phase shift signals. GCI have been utilized for optical label free biosensing (Kozma et al., 2011), for screening molecules on extracellular vesicles (Patko et al., 2013). A good review article on waveguide interferometers have recently been published (Kozma et al., 2014). Resonant Waveguide Grating (RWG) is another example of a sensor, which employs suitable designs such as gratings to allow resonant coupling of light into a waveguide at normal incident angles, unlike SPR sensors, which have a fairly large incident angle. These gratings allow RWG sensors for real-time signal detection on larger array for example, a 384-well plate or 1536-well microtiter plates (Fang et al., 2006; Li et al., 2006). Fig. 1d shows a schematic with a colored polarized light (arrow) indicating the range of incident wavelengths, that could be applied to illuminate the waveguide at specific wavelengths for coupling and propagation along the waveguide. The resonant wavelength at which a maximum in coupling efficiency is achieved is a function of the local RI at/or near the sensor surface (Fang, 2010). Reverse symmetry (RS) is another design of an optical waveguide sensor such that the RI of the waveguide substrate (i.e. RI < 1.33) is less than the RI of the medium/buffer covering it (i.e. RI ~ 1.33). This reverse arrangement allows tuning the penetration depth permitting detection of biological matters (i.e. a bacterial of 0.5–5 μm).
Fig. 1. Type of optical biosensors. (a) Surface plasmon resonance (SPR) sensor (b) classical waveguide interferometer (WGI) or phase shift sensor (c) grating coupled interferometry (GCI) sensor (d) resonance grating waveguide (RGW) sensor (e) reverse symmetry (RS) sensor (f) metal clad waveguide (MCW) sensor.

Fig. 2. Resolution chart of different types of sensors. The blue squared area shows the limit for precise detection of biomarkers. The grey squared area shows the signal-to-noise area due to non-specific binding of analytes in the range of $K_a = (10^2-10^{11}) M^{-1}$. An ideal biosensor would be one that falls in the black dotted line range having lower detection limit at shorter assay time. SMR: suspended microchannel resonators; MC: microcantilever; NW: electrical nanowire; LFA: lateral flow assay; IFA: immunofluorescent assay; MRR: microring resonators; QCM: quartz crystal microbalance; BBA: biobarcode DNA amplification, partially adapted from (Arlett et al., 2011), Copyright 2011 with permission from Springer Nature.
with any size to be probed by the evanescent field unlike SPR sensors, where the penetration depth (∼100–150 nm) is fixed by choice of metal (Fig. 1e) (Hurváth et al., 2005). RS have been demonstrated for various biosensing applications such as using nanoporous substrate (RI ∼ 1.193) and an aqueous cover (RI ∼ 1.331) (Hurváth et al., 2002b), for biomedicine sensing of biological analytes (Hurváth et al., 2002a), as a refractometer for determining the absolute RIs of both liquids or gases (Skivesen et al., 2003). Metal clad waveguides (MCW), is another example of a sensor which utilizes low index buffer layer with appropriate thickness between the core and metallic layers as shown in Fig. 1f to attenuate and enhance TM (transverse magnetic) waves w.r.t TE (transverse electric) waves. These sensors are particularly applied for integrated optics based applications (Skivesen et al., 2007).

These strategies have helped improved the performance, stability, and sensitivity of NHA for point-of-care (POC) biosensing applications offering serum unique advantages such as: (1) real-time monitoring, (2) label-free recognition of ligands or analytes, (3) low sample volumes, (4) reusability of sensors for example, NHA sensor chips, (5) allowing testing on crude samples without purification for example, blood, urine, or serum samples, (6) repeatability, (7) ability for multiplex sensing of several analytes (Sinton et al., 2008), (8) possibility of optical integrations, or for planar integration into microfluidic setups (Escobedo, 2013), (9) the high density packing due to the efficient fabrication process enables reduced sensor chip area with increased flux improving binding and mass transport (Dahlin, 2012), and (10) better optical response due to the flexibility of NHA geometry such as hole (Gordon et al., 2004), lattice orientation (Gordon et al., 2005), and periodicity (Gordon et al., 2006). Associated with these advantages there are a few disadvantages of NHA sensors such as, high cost of micro/nano-fabrication of the sensor chips, limited throughput despite of the good fabrication, low lifetime depending on the analyte of interest (Hill, 2015), and requirement of frequent calibration due to sensor drift. Furthermore, sensor optimization is a key parameter, for example in case of miniaturized sensors, the sensitivity may not be affected but the detection limit might. In another case, the latter might not be affected but it might affect signal transduction (Dahlin, 2012). We have also presented a resolution chart comparing the performance of different sensors as shown in Fig. 2 showing the detection limit versus the analysis time.

The nanoscale dimensions of NHA has enabled development of portable and innovative healthcare devices for detection of analytes at concentrations as low as picomoles. Additionally, the tunability and flexibility of NHAs offers an opportunity to design smaller, faster, and more efficient devices. A number of review papers are available on NHAs but most of them focus on NHA based applications such as on-chip NHA sensors (Escobedo, 2013), optical properties of plasmonic biosensors (Hill, 2015), NHA as optofluidic elements (Sinton et al., 2008), NHA for chemical analysis (Masson et al., 2010), advantages and disadvantages of miniaturized sensors (Dahlin, 2012) to name a few. In the present review paper, we emphasize the principle, geometry, and the effects of various parameters involved in the design of NHAs for POC applications that has not been addressed previously. This review also highlights the sate-of-the art of the NHA fabrication techniques that has not been reported earlier. This review is broadly divided into four main sections. The first section will focus on NHA design criteria and its influence, the second section will highlight NHA fabrication techniques. The next section will emphasize on NHA based biosensors in a broader sense to design POC devices. The final section will cover the current challenges and discuss the future prospects of NHA for POC applications.

1.1. Category of NHA sensors

The design of a NHA is important as the spectral resonances can be tuned based on the design for any applications. Based on their physical structures, the nanohole sensors that have been reported can be categorized into the following 6 types: (1) through holes in a metal film supported on glass or quartz substrate (Fig. 3a), (2) through holes in the top metallic film formed in a metal-dielectric-metal sandwich film supported on glass or quartz substrate (Fig. 3b), (3) a metal film with holes with metallic nanodiscs at the bottom of the holes (straight wall structure) (Fig. 3c), (4) a metal film with holes with a metallic layer at the sidewall and bottom of the holes (tapered wall structure) (Fig. 3d), (5) nanosphere lithography mediated convex metal hole (Fig. 3e), and (6) suspended through holes on metal film supported on thin SiN film (Fig. 3f). For example, (Genet and Ebbesen, 2007) perforated holes with a period of 550 nm (letter ‘h’) and 450 nm (letter ‘v’) on Ag silver film and observed transmission of red and green colors upon illumination with white light (Fig. 3a). This is because the transmission wavelength (λmax) is proportionate to the lattice constant (p) of the NHA. Thus, when the ‘p’ of the Ag-NHA increased from 300 to 550 nm, λmax changed from 436 to 627 nm, in the visible RGB colored spectrum generating red and green colored letters ‘h’, respectively. Hence, subwavelength NHAs can be utilized as optical filters by altering ‘p’ to generate the desired transmitted color. In another similar study, (Cheng et al., 2015) established color printed structures based on plasmonic meta-surfaces with triangle shaped lattices and circular NHA fabricated on Ag-SiO2-Ag three-layered surfaces. Fig. 3b shows the highlighted imageries of MISSOURI S&T. The scanning electron micrograph (SEM) of the fabricated pattern outlined six assorted triangle shaped lattices with their correspondent colors as shown in the optical microscopy images. By selecting an appropriate lattice period and hole radii to each pixel, they were able to produce green and yellow colored athletics mark (top right). Addition of extra four distinct colors revealed that the fabricated metasurfaces supported strong absorption resonances, producing colored images with high visual contrast and consistency (‘8’ orange, ‘S’, ‘T’: magenta, pickaxe: cyan, and ‘MISSOURI’: navy blue) (bottom right). (Baca et al., 2009) proposed plasmonic crystals that consisted of Au (gold) layered plasmonic nanostructures on a polymer film as substrates meant for SERS (Surface enhanced Raman spectroscopy). The quasi-3D plasmonic crystals exhibited SERS in ∼ 105 order over a huge area with good sensitivity and spatial resolution as observed in Fig. 3c (bottom middle). These NHA based SERS substrates can open attractive applications such as portable detectors, ecological monitors, for noninvasive imaging, etc.

Color changing property of the Roman Lycurgy cup depending on the light direction from the metal-nanoparticle-optical scattering has inspired plasmonic research lately. With this idea, (Gàrta et al., 2013) designed a Lycurgy cup array (nanoLCA) setup that appeared green in transmitted illumination and changed the color to red in reflection images. Fig. 3d presents the SEM images of nanoLCA with a funnel molded cup setup immobilized with nanoparticles (NPs) in the sides (inset), and at the cross-sections (bottom right).

Established NHA platforms, now range from single-operation to entire lab-on-chip systems.

Of particular interest are, NHA based optofluidic platforms assembled by combining microfluidics and optics. (Lee et al., 2011) fabricated such systems (Fig. 3e) by spin-coating silica particles on poly-styrene (PS). The plasmon resonances can be modulated by the selection of liquids to differ the refractive indexes (RI), and by tuning geometric parameters such as size, shape, and time of incubation of the silica particles to achieve label free chemical and biomolecular sensing. NHA based sensors are strongly limited by the passive diffusion due to mass transport of analytes at the sensor surface rather than by the sensors intrinsic detection capabilities. This is profound at lower analyte concentrations, leading to longer detection times from days to weeks. To overcome this fundamental limitation, (Altug et al., 2010), merged photonics and fluidics and constructed a multilayered nano-plasmonic-nanofluidic sensor scheme that allowed 3D control of flow for efficient analyte delivery and improved response time. As shown in Fig. 3f, the NHA was suspended in a multilayered scheme connected with fluidic chambers and enabled with multiple in- and outlets on both
sides of the top/bottom chambers to actively control the fluid flow in 3D.

1.1.1. Extraordinary optical transmission

Light transmission through a hole is dependent on the diameter and depth of the hole. Mathematically (Bethe, 1944) represented transmission by:

\[ T = \frac{4(kr)^{2}}{2\pi^{2}} \text{ where, } k = \frac{2\pi}{\lambda} \text{ is the wavevector of the light with wavelength } \lambda \text{ and } r \text{ is hole radii.} \]

Since \( T \sim \frac{1}{r^4} \), the transmission efficiency decreases dramatically when \( r \ll \lambda \). In fact, the transmission becomes close to zero when light passes through sub-wavelength holes such as \( r < \lambda/4 \) (or \( d < \lambda/2, d \) is the diameter of the hole). However, Ebbesen and other groups showed that when sub-wavelength holes are made on coinages metals such as Au, Ag, and Al (Aluminum), the transmission efficiency is 3-4 order of magnitude higher than the prediction by Bethe’s equation. These EOT of light is generally attributed to surface plasmon resonance and diffraction of light wave created on the metal-dielectric interfaces of the plasmonic holes. The design criteria of NHAs are dependent on various parameters which are explained below:

1.1.1.1. Periodicity. The peak positions in a spectrum is relative to the periodicity of an array and by controlling the Azimuthal rotation, the rate of degeneration of plasmonic modes at non-zero incident angles occurring due the periodic positioning can be controlled (Gao et al., 2010). The resonance peak position of the literature data (Fig. 4b, c) follows different trend for Au and Ag based EOT substrates. In general, larger pitch produces resonance peak positions at higher wavelength (that is resonance peak of larger pitch red shifts compared to lower pitch). The relationship of resonance wavelength and pitch is as follows:

\[ \lambda_{\text{Au}}(\text{nm}) = 304.89 + 1.15p \]
\[ \lambda_{\text{Ag}}(\text{nm}) = 322.74 + 0.94p \text{ for } 200 < \lambda < 460 \text{ nm} \]
\[ \lambda_{\text{Au}}(\text{nm}) = 434.67 + 0.38p \text{ for } 500 < \lambda < 1200 \text{ nm} \]
\[ \lambda_{\text{Ag}}(\text{nm}) = 711.9 - 1.628p + 0.00258p^2 \text{ for } 300 < \lambda < 1000 \text{ nm,} \]
where \( p \) = pitch in nm

1.1.1.2. Hole/aperture diameter. Hole/Aperture design is an important parameter when designing NHAs. Larger diameter allows propagating mode of transmission to be controlled by waveguide mode resonances.
But, once diameter is less than or equal to half of the transmission spectrum, the hole permits only tunneling modes showing the interplay between surface plasmon polaritons (SPPs) and hole diameter. This can be tuned by changing the etch time in reactive-ion etching (RIE) microfabrication step (Homola, 2008). Changes in the diameter of the hole also influences the optical properties of NHA from transmission to absorption spectrum. At same wavelength ($\lambda$) holes with smaller diameters display EOT, whereas the larger ones produce Grating Coupled Attenuated Transmission (GCAT).

In order to understand the sensor performance with respect to the hole size (Ohno et al., 2016), carried out simulations at the dip in transmission as well as at the wavelength peaks. Fig. 4d shows the simulated transmission spectra of a hexagonal closed packed (hcp) NHA fabricated by self-assembled colloids. For smaller holes, the dip wavelength remained unchanged but for larger holes the transmission dip wavelength started to blue-shift. The dip and peak wavelength is directly proportional to the hole diameter due to self-polarizability of the hole. The relationship between resonance wavelength and hole diameter ($d$) over the range of 40–240 nm are as follows:

$$\lambda_{\text{peak}} = 557.9 + 2.76 d - 0.009 d^2$$

$$\lambda_{\text{dip}} = 626.6 - 3.4 d + 0.01 d^2$$

1.1.1.3. Surface roughness and thickness. Surfaces with 10–100 nm thick roughness is capable of generating electromagnetic fields of SP from the metal-dielectric interface that can either reduce the propagation of SP or produce undesirable “hot spots” (Lindquist et al., 2010; Nagpal et al., 2009). A surface is considered rough, if a plane surface has non-uniformity, or there is a height variation in this non-uniformity. Surface thickness, on the other hand, is basically the order of skin depth from the surface and is correlative with roughness. Fig. 4e, shows the variation in resonance wavelength versus thickness which is largely found within the 300 nm range, and is optimal for transmission phenomena to occur. It guides the SPPs excited by the incident light at the front boundary to couple and produce SPPs at the other boundary, when the adjacent dielectric media are alike. Besides contributing to metallic-dielectric interface, thicker films also contribute to EOT since the decay of light on the metallic films are much profound than the SPPs produced. This varies, based on shape of the array, nature of metal, and type of mode (e.g.: air mode, glass mode, silica mode, glassy carbon mode etc.). Au, Ag and Al films have been commonly used for plasmonic applications although Au films are preferred as it is chemically inert, non-cytotoxic, amenable, and less rough (Love et al., 2005). Alternatively, Ag and Al films have low optical absorption losses and low cost, but it readily oxidizes in air or water (Im et al., 2011).

1.1.1.4. Depth of hole/aperture. The generated transmission spectrum is directly related with the hole/aperture depth. Literature suggests that the metallic layers punctured with an array of periodic nanoholes exhibits EOT at particular wavelengths (Degiron et al., 2002). This enhanced transmission is a result of SP channeling from one side of the hole to the other side.

1.1.1.5. Order of the hole. The order of the hole is another important parameter. For example, for a single aperture in a flat film, the highly intense electromagnetic fields (known as “hot-spots”) can result in two consequences that effect the transmission. (1) An increase in the hole’s diameter such that the cutoff wavelength restricts the propagative modes inside the holes, and (2) field enhancement on the rim, where the excited LSPRs upsurges the transmission. Degiron and Ebbesen claimed that the role of localized modes in EOT was insignificant in comparison to the propagating SPPs (Nakamoto et al., 2011). This is because NHAs provides SPPs free space to scatter light from the neighboring holes leading to EOT whereas in isolated holes the SPPs
from one hole are mostly dissipated and cannot contribute to enhanced transmission.

1.1.2. Single hole versus multiple hole arrays

According to Maxwell’s 1st order equations the scattering of light is controlled by the shape and permittivity of the object. Hence, NHAs designed with single and multiple holes should display different behavior. This behavior is dependent on two parameters: the lattice periodicity and the polarizability. The periodicity governs the amount of induced dipoles required to produce total reflection from small sized particles or total transmission by small holes. Polarizability on the other hand is directly proportional to the particle or hole radii. Several types of NHAs have been fabricated so far. For example, (Rajasekharan et al., 2014) presented a submicron plasmonic color of NHAs have been fabricated so far. For example, (Rajasekharan et al., 2014) presented a submicron plasmonic color filters with sub-RGB filters (size of red colored rectangle is 1.95 μm × 5 μm (RGB)), adapted from (Rajasekharan et al., 2014), Copyright 2014 with permission from The Nature Publishing Group. (c) Schematic of the multispectral device, which consists of 4 × 4 block array, where each block contains a 3 × 3 grid of cells (each cell represents one NHA with a unique periodicity p (in nm)); a colored image of the device is shown (bottom left); Optical transmission spectra for nine NHAs within a single block (right), adapted from (Najiminaini et al., 2013a), Copyright 2013 with permission from Springer Nature. (d) Optical microscope images of aluminum colored filters with circular, square and triangular shaped holes arranged on hexagonal and square lattice are shown, adapted from (Inoue et al., 2011), Copyright 2011 with permission from American Institute of Physics.

1.1.3. Periodic hole structure versus random hole structures

The diffraction pattern in NHAs can be manipulated by fabricating either periodic/ordered or non-periodic/random hole arrangement to harness either far-field or near-field interferences. The performance of such spatial arrangement depends on two parameters; period p and wavelength λ. Periodic nanostructures support Bragg scattering, super-radiance, plasmonic band gaps, Fano resonances, etc. whereas random nanostructures supports Anderson localization, and anomalous transmission enhancement (Conley et al., 2014; Segev et al., 2013).

The absorption enhancement in random holes is more than periodic holes due to multiple scattering (Bao and Ruan, 2010), and better matching of spectra using leaky mode resonance (Donnelly et al., 2014; Cao et al., 2009). Literature suggests that for nanostructures with hole size equal to its wavelength, the light absorption enhancement can exceed the conventional Yablonovitch 4π limit offering larger space to optical design at the nanoscale (Yu et al., 2010). Many photon management schemes have been proposed to enhance light trapping in semiconductor thin films, such as one-dimensional grating (Zeng et al., 2006), photonic crystal (Basu Mallick et al., 2012), plasmonics (Catchpole and Polman, 2008; Ferry et al., 2010), random textured surface (Rockstuhl et al., 2008), textured transparent electrode (Millier et al., 2004), nanowires or nanoholes (Tsakalakos et al., 2007; Hu and Chen, 2007, 2010). Fig. 5c (bottom left) shows the fabricated multispectral device captured with a colored CCD camera. The difference in transmission resonances and minima from each cell resulted in variation of visible color as seen in Fig. 5c (right). Multispectral imaging via periodic NHAs could be a gateway for wide range POC applications from remote sensing to biomedical imaging as they do not suffer from performance degradation or crosstalk effects in spectral filtering when they are closely packed on a single platform (Vasefi et al., 2011).
1.1.4. Short-range ordered versus periodic nanoholes

Periodic nanoholes are designed in continuous metallic substrates while short-range ordered nanoholes can be formed using metallic NPs. Both are electrically conductive, have similar electromagnetic field localization, and bulk sensitivities. The value of momentum component of light in metal films perforated with square array of nanoholes is defined as:

\[ k_{\text{photon}} = \frac{2\pi}{p} (\sqrt{i^2 + j^2}) \]  

(1)

where, \( p \) is lattice constant (i.e., distance between nanoholes), \( i' \) and \( j' \) are non-zero integer numbers representing scattering orders from 2D array.

For an opaque metal film perforated with periodic array of nanoholes, wavelength of light that can be transmitted is defined in Eq. (2) (Ebbesen et al., 1998b; Genet and Ebbesen, 2007).

\[ \lambda_{\text{peak}} = \frac{p}{\sqrt{i^2 + j^2}} \left( \frac{\varepsilon_m R^2}{\varepsilon_m + \varepsilon_d} \right) \]  

(2)

where, \( \varepsilon_m \) is dielectric function of metal.

1.1.5. Square verses triangular lattice

Different lattice structure will allow transmission to specific plasmonic resonance wavelength. Both Square and Triangular lattices based plasmonic structures have been constructed for color generation (Prasad et al., 2016; Betancourt and del Río, 2006). The transmission wavelength for square lattice NHAs is defined as:

\[ \lambda_{\text{max}} = \frac{p}{\sqrt{i^2 + j^2}} \left( \frac{\varepsilon_m R^2}{\varepsilon_m + \varepsilon_d} \right) \]  

(3)

where, \( \varepsilon_d \) is the permittivity for dielectric medium.

Eq. (3), neglects presence of holes, scattering loss, and Fano form of interactions that causes a red-shift resulting in a small transmission peak wavelength than predicted. Also, \( \lambda_{\text{max}} \) is proportional to \( p \), showing that it can act as optical filters by simply selecting transmitted color and adjusting the lattice constant. In comparison to square-lattice, triangular-lattice NHAs have better color filtering performances. The \( \lambda_{\text{max}} \) for triangular-lattice ‘\( p \)’ is:

\[ \lambda_{\text{max}} = \frac{p}{\sqrt{i^2 + j^2}} \left( \frac{\varepsilon_m R^2}{\varepsilon_m + \varepsilon_d} \right) \]  

(4)

1.1.5.1. Hole shape. The hole shape guides the spectral position of resonances creating high local fields, and the holes arranged in periodic arrays guides the band gaps (Van der Molen et al., 2005). As mentioned previously, the hole shapes could be circular, triangular, squared, or rectangular. Also, (Inoue et al., 2011) explained how shape guided the spectrum by fabricating Al built RGB-colored filters with subwavelength-holes as shown in Fig. 5d. The generation of diverse color is due to the difference in hole shape and periodicity as shown in the corresponding SEM images (inset). Such square and triangular NHA architectures have attracted huge interest for fabrication of several POC biosensors (Sulka et al., 2017; Joannopoulos et al., 1995; Gadot et al., 1997; Kyoitani et al., 1996; Che et al., 1998).

1.1.5.2. Optical performance. NHA optical performance is defined by three important factors: sensitivity, resolution, and limit of detection (LOD).

1.1.5.3. Sensitivity. It is the variation in sensor output to RI, and is achieved by either changing the physical dimensions to control the resonance wavelength, or by tuning the penetration depth by changing the angle of incidence. Sensitivity is measured in terms of RIU (Refractive Index Unit).

\[ S = \frac{\Delta \lambda}{\Delta n} = \frac{\lambda_{\max} \varepsilon_m (\varepsilon_m + \varepsilon_d)^2}{\sqrt{i^2 + j^2}} \left( \frac{\varepsilon_m R^2}{\varepsilon_m + \varepsilon_d} \right)^3 \]  

(5)

where, \( S \) is sensitivity and \( n^2 \) is the RI of the solution in contact to the NHA surface.

1.1.5.4. Resolution. It is described as the least minimal detectable variation in bulk RI which is mostly limited by the noise of the instrumentation.

\[ R = \Delta n \cdot \Delta \lambda \]  

(6)

where, \( R \) is the resolution.

1.1.5.5. Limit of detection. In contrast to bulk variation in resolution, LOD is described as the minimum amount of any molecule or substance that can be detected (Balasa, 2015).

2. Principle of operation of SPR sensors

SPR biosensors are based on the principle of interaction of SP at the metal-dielectric medium with respect to a change in the RI. Based on the relationship between RI, dielectric medium, and propagation of plasmons, SPR sensors can be categorized into 3 classes: (1) Angular mode, in which the angle of incidence is varied to excite SPs and the shift in the wavelength is measured (for example, prism-coupling based SPR sensors); (2) Wavelength mode, where a polychromatic optical source is used for plasmon excitation and a dip in transmitted or reflected light wavelength is observed, (3) Intensity mode, where, both incident angle and wavelength are constant and the transmitted or reflected light intensity is measured, and (4) Phase mode, where shift in phase of incident light is measured (Homola, 2006). Furthermore, since organic components have bulk RI ~ 1.5, higher than water (~ 1.33) and air (~ 1), bio-recognition molecules can be used to improve the sensitivity (by achieving RI matching), for example, employing binding of avidin-biotin, biotin–biotin labelled proteins, biotin-nucleic acids on the plasmonic surfaces (Wijaya et al., 2011).

2.1. Effect of reflection and transmission modes in detecting SPR for POC applications

There are two modes of SPR detection namely, Reflection and Transmission modes. In reflection mode illuminated light and detector are on the same side whereas for transmission mode light and detectors are on the opposite sides of the sensor. SP excitation is generally robust at the side of incident light projection making reflection mode much superior. The transmission mode performance can be enhanced by use of cardioid condensers, or by using detectors with small numerical aperture, or by using optical filters to remove the undesirable scattered light and improve the signal-to-noise (S/N) proportion (Wu et al., 2012b). NHAs can be designed based on the demand of optical power flow for various applications (Wang et al., 2009; Yu et al., 2004; Wu et al., 2012a).

3. Fabrication techniques of NHA structures for plasmonics and nanophotonics

In this section, we will focus on the fabrication techniques of periodic NHA because of its performance and its potential to transform the field of plasmonics and nanophotonics in a deterministic manner. Usually, production of NHAs requires a series of micro/nanomachining processes including deposition/coating, etching, micromachining and nanofabrication steps. Development of fabrication protocols and selection of fabrication techniques for nanohole array structures for plasmonics and nanophotonics require consideration of the design of lateral and vertical structures and the materials involved in the structures in addition to the target applications. While the selection of
specific fabrication techniques to produce nanoscale structures mainly depends on the design of lateral structures, other process steps such as etching and deposition and their process sequences determine the design of vertical structures. Following this classification, we first discuss the overall process sequences and fabrication techniques to produce vertical structures, i.e. etching and deposition techniques, which will be followed by various nanofabrication techniques to produce lateral NHA structures.

### 3.1. Overall process scheme

The metallic NHA can be fabricated in three ways as shown in Fig. 6. In the first way (Fig. 6a), a metal layer is deposited prior to the nanopatterning step, and the metallic NHA structures are achieved through subsequent etching of the metal layer using nanopatterned resist as an etch mask (Cetin et al., 2015, 2018; Kim et al., 2012b; Zeng et al., 2017). In the second way (Fig. 6b), the nanopatterning step is followed by metal deposition, and the formation of a NHA is achieved through a lift-off process (Grigorescu and Hagen, 2009). The sequence of metal deposition is dependent on the materials compatibility with metal etching or lift-off process. These two-basic fabrication processes have been combined to produce more complex NHA devices. An example is NHA in a suspended membrane (Escobedo, 2013; Kumar et al., 2014), where fluidic access is available from both sides of the membrane. In the third way, a freestanding membrane, mostly a Si$_3$N$_4$ (silicon nitride), undergoes a combination of wet chemical etching of Si and photolithography, followed by nanopatterning and metal deposition as shown in (Fig. 6c).

### 3.2. Metal etching and lift-off

**Metal etching** process can be divided into wet and dry etching. In wet etching, a wet chemical etchant preferentially dissolves a specific metal layer over the patterned resist in an extremely selective manner. Numerous etchants specific to different metals are readily available with well-developed protocols. Wet etching has been widely used for micromachining process, but its utilization in nanofabrication was rather limited due to isotropic nature of this etching technique which produces an undercut of the polymer resist by lateral etching under the patterned mask. Also, control of an accurate etching time is hard to achieve. Au etch rates are about 680 nm/min and 660 nm/min (diluted with 3 parts 37% HCl + 1 part 70% HNO$_3$ + 2 parts H$_2$O) and AU-5 etchant (5% I$_2$ + 10% KI + 85% H$_2$O, Cyantek), respectively. Ag etch rates are about 450 nm/min and 600 nm/min with CR-7 etchant (9% (NH$_4$)$_2$Ce(NO$_3$)$_6$ + 6% HClO$_4$ + 85% H$_2$O, Cyantek) and a piranha solution (50 parts 96% H$_2$SO$_4$ + 1 parts 30% H$_2$O$_2$), respectively (Williams et al., 2003). Unlike wet etching, dry etching is an anisotropic physical etch process with high selectivity. Dry etching or RIE requires a plasma assisted system, where reactive species such as radicals and ions are generated and directed towards the substrate to be etched. A representative RIE system for anisotropic etching is the inductively coupled plasma (ICP)-RIE. The etching selectivity of different materials and anisotropy in the etching direction has made ICP-RIE most suited not only for micromachining process but also for fabricating various nanostructures (Williams et al., 2003; Williams and Muller, 1996). Si, SiO$_2$, and Ge, are the common semiconducting materials used for RIE, however no well-defined etch processes exist for metals, like Au and Ag which are usually used for plasmonics and nanophotonics. For RIE of metals, chlorine (Cl$_2$, CCl$_4$ or BCl$_3$) or bromine (HBr) etch chemistries are often used (Green, 2014; Ranade et al., 1993). Chlorine etching leads to formation of metal halides from metals like Au, Ag, Pt, Fe, Ni, and Co and usually leads to a large amount of redeposition. The redeposition can be reduced by using a heated chuck to maintain the substrate at an elevated temperature. Another method to reduce redeposition, as shown by (Ranade et al., 1993) is to use a gas mixture of CF$_4$ and CCl$_4$.

In contrast to etching process, the **lift-off process** is used to remove the patterned resist layer by a solvent after a metal deposition. Deposition of metals at sidewalls of the resist pattern and the shadowing effect by the resist pattern result in edge roughness in the remaining metal patterns on the substrate. Also, lifted-off metals in the bath can be redeposited on the sample surfaces during the entire process (Lindquist et al., 2012). A common strategy to facilitate the lift-off process is the use of multi-layered resists where a lift-off resist was coated first followed by the coating of the resist for patterning. During the development process following patterning, undercuts are produced under the patterned resist, which prevents sidewall deposition during metal deposition and facilitate the lift-off process (Alayo et al., 2015; Barnett et al., 2017; Calafiore et al., 2014; Lyman et al., 1981; Wang et al., 2015). The lift-off process is tedious for nanoscale patterns required for plasmonic applications, due to the decreased thickness of the resist layer. Because of those issues, a technique which does not require the lift-off process after a metal deposition has been lately developed by (Yanik et al., 2011a, 2010a, 2010b). The key of this technique is to use a freestanding Si$_3$N$_4$ membrane having NHA structures as a stencil mask for metal deposition, so that the metal to be deposited flows out through the NHA and only remains on the surface during the metal deposition. However, diffusion of metal atoms underneath the stencil mask makes it difficult to accurately control the nanopattern dimension. Also, re-usability of the stencil mask for multiple metal deposition can be an issue in expanding this technique as a versatile fabrication method for nanopatterning.

A thin adhesion layer is usually needed to deposit between the substrate and the metal layer, in particular for the Au layer. The typical
adhesion layer used is 1–10 nm thick Ti (Cetin et al., 2018; Coskun et al., 2014; Huang et al., 2013; Kee et al., 2013; Najiminaini et al., 2011; Park et al., 2018; Sannomiya et al., 2011; Soler et al., 2017; Yanik et al., 2011a, 2010a), Cr (Ferreira et al., 2009; Im et al., 2010; Lesuffleur et al., 2008a; Najiminaini et al., 2011; Wang et al., 2014; Xiong et al., 2016; Xue et al., 2017), or NiCr (Sharpe et al., 2008). It should also be mentioned that the composition and depth of the adhesion layer affect the optical behavior significantly (Najiminaini et al., 2011).

3.3. Patterning techniques in resist

**Deep ultra-violet (UV) lithography (DUVL)** is based on a conventional photolithography. In the photolithography, PR spin-coated on substrates is exposed to UV light through a photomask which is patterned with metals such as chromium (Cr). If PR is positive tone, the area exposed with UV light through the Cr pattern will be dissolved during developing, while if PR is negative tone, PR will remain during developing. Normally, the photolithography is limited by a wavelength of UV light, i.e. the smallest pattern size theoretically equals to the square root of the light source (e.g. wavelength) if a photomask is in contact with a substrate (Seisyan, 2011).

To overcome this limitation, short wave length of UV such as DUV or excimer laser (e.g. KrF, ArF, and F2) with a projection method can be used (Seisyan, 2011). In the projection method, a photomask does not contact with a substrate and several objective lenses will be positioned between a photomask and a substrate. In this case, the smallest pattern size follows by $\frac{1}{2\rho_n}$, where $\rho_n$ is the factor depending on the degree of coherence and the imaging method, $\lambda$ is the wavelength of the actinic radiation, and $NA$ is the numerical aperture of the lens (Seisyan, 2011). This technique allows to fabricate NHA structures on wafer scaled substrates with high uniformity and reproducibility. Indeed, DUVL was successfully used to make NHA having 200 nm diameter and 600 nm period on Si$_3$N$_4$/Si substrates (Coskun et al., 2014; Huang et al., 2013; Kee et al., 2013; Soler et al., 2017), and glass substrates (Cetin et al., 2018). However, this technique requires sophisticated objective lens and nanoscale precise alignment system as well as an expensive photomask supporting nanoscale patterns.

**Interference lithography (IL)** allows fabrication of large area NHAs without requiring any photomasks or advanced objective lens systems (Lu and Lipson, 2010; Wang et al., 2018a). Two or more coherent lights incident from varied directions form interference patterns in a photoresist layer. The period of the interference patterns is determined following $2\pi/\lambda$, where $\lambda$ is the wavelength of light source, $N$ is RI of the photoresist, and $\theta$ is the angle at which the two incident lights interfere (Lu and Lipson, 2010). Thus, the period of the interference patterns can be controlled by either the light wavelength used, the angle of incident beams, or the combination thereof. UV lasers with various wavelengths (365 and 355 nm, Ar laser at 244 nm, and Nd: YAG laser at 213 nm) have been used to create interference patterns. The minimal feature size possible with those wavelengths are 182, 178, 122 and 96 nm, respectively (Seo et al., 2014). The area of an interference pattern is limited by the diameter of cohesive illuminating beams. Fabrication of an Al polarizer with uniform interference patterns of 3 cm x 3 cm are has been reported via IL and subsequent nanoimprint lithography (Kim et al., 2004). Even smaller periodic patterns are possible by combining IL with H$_2$O immersion and/or non-linear frequency multiplication techniques (Lu and Lipson, 2010). However, this reduction in the pattern size is sacrificed by reduction of the area of cohesive beams, chiefly the pattern area. Meanwhile, as shown in Fig. 7a, (Henzie et al., 2007) used IL to produce PDMS molds for subsequent phase-shifting photolithography (PSP) to produce large area of NHAs with 100 nm diameter, 400 nm period and 400 nm height. They demonstrated fabrication of hundreds of PDMS molds to be used as a mask for PSP from a single IL master without any defects.

Electron beam lithography (EBL) is a direct writing method without need of a photomask such that smaller features can be achieved in comparison to optical lithography and IL due to the much smaller wavelength of electron beam (Sanchez and Sobolov, 2010). In EBL, nanoscale patterns are written on an e-beam sensitive resist coated on a target substrate (Guo, 2007). Poly(methyl methacrylate) (PMMA) and hydrogen silsesquioxane (HSQ) are representative positive and negative resists, respectively, for EBL process. For high resolution EBL, a very thin layer of resist is required to minimize the electron scattering effect in the resist (Grigorescu and Hagen, 2009). Sub-10 nm features have been routinely achieved with EBL. Correct selection of resist/developer and careful post exposure fabrication process is crucial for fabricating well-defined NHAs. As shown in Fig. 7b, EBL facilitated NHA have been fabricated using different materials such as Si$_3$N$_4$/Si (Yanik et al., 2011a, 2010b), Si$_3$N$_4$/glass (Cetin et al., 2015). Pyrex (Najiminaini et al., 2013b). The scalability of EBL is limited because it is generally used as photomasks, striped-templates, or as nano-imprinted molds.

**Nanoimprint lithography (NIL)** is a replication technique utilizing conformal contact between a mold and a resist-coated substrate. This technique does not require any light or high energy beam sources and therefore, the spatial resolution is not limited by factors such as wave diffraction, scattering, and interference (Chou, 1996; Chou et al., 1995, 1996; Guo, 2007; Schift, 2008). There are two types of NIL namely, thermal and UV NIL. The NIL process mainly consists of two steps: molding and demolding. In the molding process, a thermoplastic resist above its glass transition temperature ($T_g$) flows into the cavities of the mold. The resist becomes hardened either by reducing the temperature below its $T_g$ or by UV-curing for both thermal and UV NIL, respectively. In the demolding step, the mold is detached from the molded resist and reused for the next molding cycle. Both thermal (Im et al., 2011; Kumar et al., 2014; Lee et al., 2015; Martinez-Perdiguero et al., 2012; Martinez-Perdiguero et al., 2013; Qi et al., 2018), and UV-NIL (Chen et al., 2009; Nakamoto et al., 2012; Verschuuren et al., 2015; Wong et al., 2013) have been used to produce large area metallic structures for SP biosensing applications. For most cases, nanostructures were formed in a thin polymeric resist layer by NIL. However, (Yu et al., 2012) demonstrated corrugated Au structures formed via direct nano-imprinting into thin Au film coated on PC (photonic crystal) substrates. NHA fabricated in Si via NIL and RIE can also be used as a template for metal deposition and reprocessed to multiple identical NHAs (Im et al., 2011).

There are a couple of issues to consider in utilizing NIL for fabricating NHA structures, which include mold fabrication and presence of a residual layer after imprinting. Imprint molds are usually made of inorganic materials with high Young’s modulus such as Si and quartz to achieve high replication fidelity upon conformal contact and created using high-end nano fabrication tools such as EBL and FIB. Fig. 7c, shows UV-curable resin used as NIL mold to replace hard molds, taking advantage of its lower Young’s modulus and thus lower demolding force (Choi et al., 2018; Jia et al., 2018; Wu et al., 2011). The expensive original Si master mold is only needed when the resin mold is damaged and needs to be remade. Thus, the lifetime of the Si master mold can be significantly prolonged. The other issue is that for most NIL process a thin residual layer remains and thus a subsequent RIE process is usually followed to remove the residual resist layer. Recently, (Lee and Jung, 2005) developed an NIL process with non-residual layer using benzyl-methacrylate monomer solution as a NIL resist, to avoid the additional etching process.

3.4. Resist-free patterning techniques

**Focused ion beam lithography (FIBL)** is a resist-free process of directly writing nanoscale patterns on target substrates by focused ion beams with higher energy (Sobolov and Sanchez, 2012). Because ions are heavier than electrons, FIBL is less affected by scattering. Unlike other nanofabrication techniques, this technique allows fabrication of
3D structures with a single process run. Generally, the minimum feature size that can be achieved by FIBL is typically 20–30 nm, which is determined by the combination of ion source selection, scan method selection, beam parameter selection, and processing parameter selection (Kim et al., 2012a). Ga is used as a common ion source for FIBL because it provides a low melting temperature, volatility, and vapor pressure (Lindquist et al., 2012). Instead Ga+ ions, He+ ion beams have been used for drilling 5 nm holes in a 100 nm thick Au foil (Ananth et al., 2011). The typical nanohole size obtained using Ga+ ions is limited to ~ 50 nm. FIBL is a tool to fabricate isolated holes in metallic substrates, but it cannot be applied to fabricate isolated metallic nanoparticles and assemblies which are commonly adopted in most applications. To overcome this limitation, (Chen et al., 2016) developed an approach of defining plasmonic isolated nanoparticle structures using FIB milling and removing the surrounding metal via a simple “sketch and peel” strategy as shown in Fig. 8a.

Colloidal lithography (CL) is a resist-free nanofabrication technique combining bottom-up and top-down approaches (Wang et al., 2018b). CL fabrication involves forming a monolayer with synthesized
micro- and nanoparticles on a target substrate. Although, this technique is cost-effective it is limited due to tedious particle patterning step (Niu et al., 2013; Velev and Gupta, 2009; Wood, 2007; Zhang and Wang, 2009). When a closed packed monolayer is formed, the substrate layer is exposed to RIE to reduce colloidal sizes, which is followed by metal deposition. This leads to the formation of hexagonal NHAs on a metal layer. In another study, (Xiong et al., 2016) successfully controlled O2 plasma conditions to reduce the size of PS particles from 140 nm to 50 nm, fabricating NHA structures having different pore sizes. The RIE step to reduce the colloidal size is needed only when nanoparticles are closed packed. For example, Fig. 8b shows NHA with a geometric gradient obtained by inclined RIE process and used as a fine plasmonic “library” (FPL) (Xue et al., 2017).

3.5. Other alternative techniques

Microcontact printing (µCP) and template stripping (TS) are high-throughput and simple techniques to fabricate NHA structures (Candeloro et al., 2017; Im et al., 2011; Lee et al., 2012; Lin et al., 2017; Xia and Whitesides, 1997). Fabricating a template made of PDMS or other materials is an essential prerequisite for these two techniques. Then, a metal deposition process is carried out to form a thin metal layer on the template surface, followed by transferring the metal layer to a target substrate. For example, (Xia and Whitesides, 1997) prepared a PDMS template by replication from a Si matter mold. The PDMS template was deposited with Au film and then the film was transferred to a target substrate treated by self-assembled monolayers (SAMs). In this case, the PDMS template was easily deformed during the printing process due to its low Young’s modulus and therefore the size of the
<table>
<thead>
<tr>
<th>Method of patterning</th>
<th>Ref.</th>
<th>Type</th>
<th>Metal</th>
<th>Nanopore diameter/period</th>
<th>Etching/lift-off or other details</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep ultra-violet (UV) lithography (DUVL)</td>
<td>(Coskun et al., 2014)</td>
<td>SiN suspended membrane</td>
<td>Au/Ti (125 nm/5 nm)</td>
<td>200 nm/600 nm</td>
<td>RIE (SF₆ + Ar) for etching SiN</td>
<td>Real-time biomolecular interactions of protein A/G with immunoglobulin G (IgG) antibody</td>
</tr>
<tr>
<td></td>
<td>(Juang et al., 2013)</td>
<td>SiN suspended membrane</td>
<td>Au/Ti (125 nm/5 nm)</td>
<td>200 nm/600 nm</td>
<td>RIE (SF₆ + Ar) for etching SiN</td>
<td>Ultrasensitive, rapid, quantitative and label-free detection of virus like analytes</td>
</tr>
<tr>
<td></td>
<td>(Lee et al., 2010)</td>
<td>SiN/glass or glass substrate</td>
<td>Au/Cr (120 nm/5 nm)</td>
<td>200 nm/600 nm</td>
<td>Lift-off</td>
<td>Monitoring growth of bacteria and antibiotic susceptibility test</td>
</tr>
<tr>
<td></td>
<td>(Najiminaini et al., 2013b)</td>
<td>SiO₂/Al (200 nm/150 nm)</td>
<td>Al (50 nm)</td>
<td>163–218 nm/260–342 nm</td>
<td>RIE (O₂ plasma to remove residual layers, and lift-off)</td>
<td>One-step simultaneous detection of Chlamydia Trachomatis and Neisseria Gonorrhoeae in urine</td>
</tr>
<tr>
<td></td>
<td>(Chen et al., 2012a)</td>
<td>glass substrate</td>
<td>SiO₂/Al (100–200 nm/150 nm)</td>
<td>80–110 nm/250–420 nm</td>
<td>RIE (SiCl₂) for etching Al</td>
<td>Surface plasmon based protein sensor</td>
</tr>
<tr>
<td></td>
<td>(Inoue et al., 2011)</td>
<td>glass substrate</td>
<td>SiO₂/Al (100 nm/150 nm)</td>
<td>150–210 nm/300–420 nm</td>
<td>RIE (CHF₃ and Cl₂ plasma to remove residual layers, and lift-off)</td>
<td>Surface plasmon based protein sensor</td>
</tr>
<tr>
<td></td>
<td>(Chen et al., 2012b)</td>
<td>glass substrate</td>
<td>SiO₂/Al (200 nm/150 nm)</td>
<td>140–230 nm/250–430 nm</td>
<td>RIE (SiCl₂) for etching Al</td>
<td>Real-time label-free protein absorption monitoring</td>
</tr>
<tr>
<td></td>
<td>(Kumar et al., 2014)</td>
<td>SiN suspended membrane</td>
<td>Ag/SiN (200 nm/200 nm)</td>
<td>200 nm/500 nm</td>
<td>Thermal-NL with unknown mold and dry etching for SiN</td>
<td>Surface plasmon based protein sensor, used to detect and identify virus-like particles</td>
</tr>
<tr>
<td></td>
<td>(Martinez-Perdiguero et al., 2012)</td>
<td>glass substrate</td>
<td>Au/Ti (60 nm/5 nm)</td>
<td>100–250 nm/450–800 nm</td>
<td>Thermal-NL with a Si stamp, O₂ plasma to remove residual layers, and lift-off</td>
<td>Real-time label-free protein absorption monitoring</td>
</tr>
<tr>
<td></td>
<td>(Martinez-Perdiguero et al., 2013)</td>
<td>glass substrate</td>
<td>Au/Ti (50 nm/5 nm)</td>
<td>100–250 nm/450–800 nm</td>
<td>Thermal-NL with a Si stamp, O₂ plasma to remove residual layers, and lift-off</td>
<td></td>
</tr>
<tr>
<td>Method of patterning</td>
<td>Ref.</td>
<td>Type</td>
<td>Metal</td>
<td>Nanopore diameter/period</td>
<td>Etching/lift-off or other details</td>
<td>Application</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>--------------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>----------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>(Qi et al., 2018)</td>
<td>polymer substrate</td>
<td>Au (65 nm)</td>
<td>275 nm/500 nm</td>
<td>Thermal-NIL with a Si stamp</td>
<td>Biochemical sensing</td>
<td></td>
</tr>
<tr>
<td>(Chen et al., 2009)</td>
<td>glass substrate</td>
<td>Au/Cr (50 nm/5 nm)</td>
<td>100-250 nm/200-500 nm</td>
<td>UV-NIL with a h-PDMS stamp and RIE (O₂/CHF₃, SF₆, and O₂) for etching</td>
<td>Biochemical sensing</td>
<td></td>
</tr>
<tr>
<td>(Nakamoto et al., 2012)</td>
<td>PET substrate</td>
<td>Au (100 nm)</td>
<td>300 nm/600 nm</td>
<td>Soft-NIL with a s-PDMS stamp and HF etching to remove sol-gel pillar</td>
<td>Chemical sensing, hydrogen peroxide from 10 to 250 μM</td>
<td></td>
</tr>
<tr>
<td>(Verschuuren et al., 2015)</td>
<td>glass substrate</td>
<td>Au (130 nm)</td>
<td>180 nm/780 nm</td>
<td>UV-NIL with a Ni mold and lift-off</td>
<td>Test optical properties</td>
<td></td>
</tr>
<tr>
<td>(Wong et al., 2013)</td>
<td>glass substrate</td>
<td>Au/Cr (50 nm/n/a)</td>
<td>140 nm/400 nm</td>
<td>Fluorescence enhanced detection of biomarkers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Brolo et al., 2004)</td>
<td>glass substrate</td>
<td>Au (100 nm)</td>
<td>200 nm/510-590 nm</td>
<td>Ga⁺ ion</td>
<td>SP-based sensors for the adsorption of biomolecules</td>
<td></td>
</tr>
<tr>
<td>(Yokogawa et al., 2012)</td>
<td>quartz substrate</td>
<td>Al (150 nm)</td>
<td>80-280 nm/220-500 nm</td>
<td>–</td>
<td>Plasmonic color filters for CMOS image sensor</td>
<td></td>
</tr>
<tr>
<td>(Menezes et al., 2012)</td>
<td>glass substrate</td>
<td>Au (100 nm)</td>
<td>305 nm ± 20 nm/700 nm ± 3 nm</td>
<td>SU8 as an adhesive layer between glass and Au</td>
<td>Comparison of plasmonic arrays of holes recorded by IL and FIBL</td>
<td></td>
</tr>
<tr>
<td>(Balaur et al., 2016)</td>
<td>quartz substrate</td>
<td>Ag/Ge (150 nm/3 nm)</td>
<td>160 nm length &amp; 40 nm width/X, 280 nm &amp; Y, 300-480 nm</td>
<td>Quartz, pattern is not a circular shape.</td>
<td>Nanoscale plasmonic pixels</td>
<td></td>
</tr>
<tr>
<td>(De Leebeeck et al., 2007)</td>
<td>glass substrate</td>
<td>Au (100 nm)</td>
<td>150 nm/350-850 nm</td>
<td>Ga⁺ ion</td>
<td>SPR-based chemical and biological sensors.</td>
<td></td>
</tr>
<tr>
<td>(Ferreira et al., 2009)</td>
<td>glass substrate</td>
<td>Au/Cr (100 nm/5 nm)</td>
<td>170 nm/500 nm</td>
<td>50 nm thick SiO₂ nanolayer on these 100 nm-thick gold films on glass</td>
<td>Attomolar protein detection</td>
<td></td>
</tr>
<tr>
<td>(Niu et al., 2013)</td>
<td>glass substrate</td>
<td>Au (20-50 nm)</td>
<td>300 nm diameter hexagonal closely packed from 330 nm PS varied 230-626 nm hexagonal lattice from 507 to 685 nm PS</td>
<td>RIE etching to decrease particle sizes after tilting</td>
<td>The detection of specific biomolecular interactions as a fine plasmonic “library” (FPL)</td>
<td></td>
</tr>
<tr>
<td>(Xue et al., 2017)</td>
<td>quartz substrate</td>
<td>Ag/Cr (140 nm/2 nm)</td>
<td>50-140 nm diameter/non-periodic</td>
<td>RIE etching to decrease particle sizes</td>
<td>Protein detection</td>
<td></td>
</tr>
<tr>
<td>(Xiong et al., 2016)</td>
<td>glass substrate</td>
<td>Au/Cr (30 nm/1 nm)</td>
<td>100-300 nm/400 nm</td>
<td>Copy from a Si template with NOA63 as an adhesive layer</td>
<td>SERS devices for biosensing</td>
<td></td>
</tr>
<tr>
<td>(Candeloro et al., 2017)</td>
<td>glass substrate</td>
<td>Au (50 nm)</td>
<td>50-140 nm diameter/non-periodic</td>
<td>Copy from a Si template with NOA61 as an adhesive layer</td>
<td>SPR bioimaging and plasmonics</td>
<td></td>
</tr>
<tr>
<td>(Im et al., 2011)</td>
<td>glass substrate</td>
<td>Ag (100 nm)</td>
<td>210 nm/500 nm</td>
<td>Copy from a Si template with NOA61 as an adhesive layer</td>
<td>BSA and anti-BSA interaction experiments</td>
<td></td>
</tr>
<tr>
<td>(Lin et al., 2017)</td>
<td>PET substrate</td>
<td>Au (100 nm)</td>
<td>180 nm/600 nm</td>
<td>Copy from a Si template with a heated PC layer as an adhesive layer</td>
<td>Plasmonic color filter</td>
<td></td>
</tr>
<tr>
<td>(Hwang et al., 2017)</td>
<td>PET substrate</td>
<td>Al (30 nm)</td>
<td>100 nm/280-430 nm</td>
<td>Copy from a PET template (copied from a Si mod) with a mr-APS-1 as an adhesive layer</td>
<td>Plasmonic color filter</td>
<td></td>
</tr>
</tbody>
</table>
Further, typical detection of exosomes requires considerable purification of the sample and labeling of the samples. Hence, NHA based detection methods provide a label-free and high throughput approach to probe exosomal markers for clinical and functional proteomic applications. In another study, (Yanik et al., 2011b) assembled a 90 × 90 µm² NHA for capturing a single layer of mouse antibody (mouse-IgG). As shown in Fig. 9b right, the resonance shift was large enough not only leading to transmission spectrum overlap but also permitting detection with the human eye without use of optical elements (Fig. 9a left). In this study, a particular design of NHA was utilized to suppress the radiative losses and utilize the Fano resonance to achieve a very high quality factor of ~ 200. The NHA was microfabricated using electron beam lithography (EBL) and interference lithography (IL) method with periodicity of ~ 600 nm to obtain very sharp Fano resonances (FWHM of ~ 4.3 nm) and high figure-of-merit of 162. (Jonsson et al., 2007) were the pioneers to address the LSPR by depositing lipid bilayers on SiO₂ NHAs for protein detection (Fig. 9c). Similarly, (Nakamoto et al., 2011) integrated Au-NHAs with microfluidic devices to detect a cell signaling protein TNF-α by performing immunoassay on the array. Fig. 9d shows calibration plot of the TNF-α with (red) or without (blue) streptavidin–Au conjugate. The inset SEM image shows Au colloids trapped on the NHA by biotin–avidin coupling. Since, the SPR response is proportionate to density of the protein layer, small molecules will result in low SPR response in comparison to larger ones. To understand this concept (Im et al., 2012) studied the binding kinetics of a single chain fragment variable (scFv) antibody molecule having the same molecular mass but different dissociation coefficients (nM). Fig. 9e shows the setup and the decrease in spectral shifts for M18 (0.19 nM), 1H (0.29 nM), 14B7 (11.7 nM), and L97A (36.7 nM) antibodies. Owing to

4. Applications

4.1. Protein detection

Protein detection is achieved by use of small proteins or bio-recognition molecules such as biotin, avidin, streptavidin, cell derived vesicles etc. for efficient ligand–receptor binding to attain high sensitivity and specificity. Recently, (Im et al., 2014) demonstrated detection of exosomes, a cell derived vesicle involved in the transport of molecular contents of cancer like cells, on NHA sensors. For this NHAs were immobilized with antibodies to detect proteins on the exosomes or in the exosome lysates. A spectral shift upon exosome binding was observed as schematically shown in Fig. 9a. The protein level of ovarian cancer specific surface markers namely; exosomal protein level expression (EpCAM) and a receptor (CD24) were higher in the ovarian cancer patients than in control groups (Fig. 9a, bottom). An advantage of this NHA sensor is that it does not require any labeling. Further, typical detection of exosomes require considerable purification and labeling of the samples. Hence, NHA based detection method provided a label-free and high throughput approach to probe exosomal proteins and serves as schematically shown in Fig. 9.
the advantages that nanoLCA offer (Plucinski et al., 2016) presented the binding of cytochrome P450, an enzyme predominantly found in cardiovascular systems with various substrates. The schematics in Fig. 9f left shows nanoLCA in gold color, proteins in red, and the cytochrome P450 in blue. Fig. 9f right shows the spectral shifts corresponding to the seven substrates at high and low concentrations. NanoLCA enabled colorimetric detection of binding of small molecules (e.g. substrate such as kI) to cytochrome P450 by allowing only one resonance plasmonic mode to transmit through the NHA. This suppression of multiple resonance modes generally associated with NHAs were due to the unique design of the nanoLCA where the conical-shaped holes were decorated with nanoparticles. Plasmonic nanoholes are the most sensitive regions of the NHA sensors. Hence, the response time of the sensor is limited by the time required for the molecules to diffuse through the small nanoholes to the sensitive regions. All the previous mentioned NHAs are based on flow-over operations. In order to improve the diffusion time of the molecules, (Gordon et al., 2008) designed a flow-through operation based NHA. Due to this microfluidic chip integrated NHA flow-through sensor, where the molecules are forced through the nanoholes, the response time of the sensor was improved by 14-times compared to the same sensor operated on flow-over principle. As a proof-of-concept experiment, (Gordon et al., 2008) observed the spectral shift upon immersion of Au-NHAs in mercaptopundecanoic acid (Fig. 9g) on the flow-through NHA. The spectrum shift (red) was observed due to variations at the dielectric medium, as predicted by Eq. (3). An additional wavelength shift peak was observed upon addition of BSA. Monitoring these minor spectral shifts can help identify unique sensing signatures making it applicable for POC devices.

4.2. Pathogen detection

Point-of-care (POC) diagnostic devices must satisfy the ASSURED criteria. ASSURED stands for Affordable, Sensitive, Specific, User-friendly, Rapid and Robust, Equipment-free, and Delivered to patient in real time detection of uropathogenic bacteria concentration is shown in Fig. 10c for diagnosis of urinary tract infections using human urine samples. The flow-through NHA sensor was able to perform the diagnosis within 35 min with a limit of detection of 100 CFU/mL, which is below the threshold limit of detection required for urinary tract infections. The intensity shift vs bacteria concentration is shown in Fig. 10c. The bar chart shows the difference in anti-UPEC and UPEC antibodies intensity shift (inset).

4.3. Cell-cell interactions

Several NHA based cell substrates have also been demonstrated. For example, (Wittenberg et al., 2012) used chemically modified NHAs using spherical lipid bilayers to measure the cell-cell interactions. The combination of spherical lipids and beads improved the S/N ratio and allowed construction of random colored arrays with single image acquisition and elimination of serial spotting. The SEM images in Fig. 11a revealed that the microwells were either occupied or empty. The spherical lipids are stable, so they have been coupled with NHAs to serve as versatile cell culture substrates for studying the cell mechanism (Wu et al., 2004; Nair et al., 2011), for understanding lipid-protein binding (Hatzakis et al., 2009), and for high-throughput drug screening of proteins coupled with receptors (Ashley et al., 2011; Roizard et al., 2011). A microfluidic-integrated NHA biosensor based on the principle of EOT for real-time analysis of live cancer cells was elucidated by (Li et al., 2017) for a period of 10 h. Fig. 11b shows the schematics of the cell and detection module and a resonance peak shift which was observed upon interactive binding of VEGF to the antibodies with a sensitivity of 145 pg/mL. In another study, (Yu et al., 2017) demonstrated label-free cell monitoring of CSH10 and HeLa cells for understanding the cell kinetics of these two different cell types. A mathematical expression was used to fit the retarded growth function and deduce the difference in adhesion dynamics. The whole cell attachment process was cross-confirmed by the cell morphology Fig. 11c (inset); cell micrographs w.r.t the mathematical fitted growth function revealed that the HeLa cells growth were slow in comparison to CSH10 cells at the same time span.

5. Current challenges and future prospects in NHA-based POC biosensing application

Although extensive studies have been reported on the fabrication, and surface modification of NHAs in quest of designing POC devices such as NHAs integrated with microfluidics (Escobedo, 2013), smartphones (Lopez et al., 2017), wearable and flexible sensors (Chen and Ming, 2012), smart materials, and metamaterials (Xu et al., 2011), its commercialization is still lacking. The major bottlenecks faced are: (1) stability of surface immobilized proteins and DNA for achieving label-free biosensing; (2) lack of multiplexing capability; and (3) lack of clinical validity. The potential solutions for these bottlenecks are addressed below:

(1) High performance biosensors with improved stability can be achieved by adopting an effective immobilization strategy such as: Bio-affinity i.e. NTA, GST, Histidine-tagged, immobilization, Thiol/Cysteine-mediated immobilization, Protein G-mediated immobilization, DNA-directed immobilization, Orientation-controlled immobilization, Lipophilic modification, or by Hydro-phobic/philic surface modification (Park and Kim, 2015). A detailed review on the immobilization strategy for biosensing applications has also been reported (Nguyen et al., 2015).

(2) Multiplexed biosensors for example, periodic NHA based microarray setups although provides a small footprint for sensing each element generally suffer from coherent noise and lack of clear interference patterns. The issues with regards to SNR can be improved by using right choice of laser beam with appropriate alignment or
Fig. 10. Pathogen detection on NHA sensor. (a) Optofluidic nanoplasmonic biosensor for virus detection, adapted from (Yanik et al., 2010b), Copyright 2010 with permission from American Chemical Society. (b) Nanopore arrays for membrane protein biosensing, adapted from (Im et al., 2010), Copyright 2010 with permission from The Royal Society of Chemistry. (c) Time-dependent uropathogenic E. coli detection assay, adapted from (Gomez-Cruz et al., 2018), Copyright 2018 with permission from Elsevier.

Fig. 11. NHA based cell culture substrates. (a) SEM of SSLBs immobilized in a microwell array, (inset) single SSLB, adapted from (Wittenberg et al., 2012), Copyright 2012 with permission from American Chemical Society. (b) Pictures of microfluidic-integrated biosensor cell culture module for real-time cytokine secretion analysis (left), Binding kinetics of the direct detection of VEGF ranging from 200 to 1000 pg/mL (right), Immunofluorescence assay for confirmation of endogenous production of VEGF in HeLa cells. VEGF was stained with Alexa Fluor 488 (green) while cell nuclei were highlighted by DAPI (blue) (inset), adapted from (Tu et al., 2017), Copyright 2017 with permission from The Nature Publishing Group. (c) Sectional view of an integrated NHA device used for single cell dynamic attachment measurement, SEM of NHA is shown in the inset; experimental results of single cell dynamic attachment process monitoring. Raw data showing the resonance peak shift, and the corresponding smoothed and fitted curve showing the binding process of a single C3H10 cell (middle), and a single HeLa cell (right). Microscopic images are also shown (inset). The microscopic images were acquired after 5, 60, 120, and 180 min of measurement, adapted from (Li et al., 2017), Copyright 2017 with permission from The Royal Society of Chemistry.
tilting, or by investing more on the fabrication (Lesuffleur et al., 2008b).

(3) Although there are serval reports published on NHA with promising biosensing applications, the transition from proof-of-concept phase to real world applications is still lacking. The possible solutions for these are: (1) More clinical trials from both healthy and diseased patients would finally be required to validate the technology for clinical analysis, (2) Biofluids selected must match the condition of the patient, (3) The assay performed must be direct and simple with a few steps, less reagents, and short assay time. (4) Issues with non-specific adsorption due to crude biofluid dilution should be resolved, and (5) The use of secondary antibodies for detection should be eliminated as it requires skilled personnel making it difficult for clinical analysis. A review addressing the issue of clinical transition of biosensors has also been reported (Masson, 2017).

Hence, the future work should be directed towards exploring and finding solutions to existing clinical problems to design better NHAs with portability, high specificity, sensitivity, good LOD, ease of fabrication, and minimal cost.

6. Conclusion

This review summarizes the current state-of-art of NHAs based POC applications. The methods of fabrications, structural and geometrical features, effects of SPR on sensing has been outlined. The design criteria of NHAs have been explained. Various design strategies of NHAs for POC applications have been presented. Finally, the current challenges and future prospects have been discussed.

Acknowledgments

This research was supported by the P41 Center for BioIndustrial Multiscale Systems for Precision Medicine (P41EB020594) from the National Institutes of Health, USA. MRG thanks the support from Louisiana Board of Regents RCS grant, USA (Contract Number: LEQSF (2017-20)-RD-A-04), and Louisiana State University start-up funds.

Declaration of interests

None.

Credit author statement

M.R.G devised the project, the main conceptual ideas and proof outline. M.R.G and S.P. were involved in the planning of the manuscript and supervised the work. A.P. took the lead in writing the manuscript. J.C. and Z.J. wrote the fabrication techniques part of the manuscript. All authors discussed the results and commented on the manuscript. None.

References


