Discovery of Amino Acid fingerprints transducing their amphoteric signatures by field-effect transistors

Naveen Kumar [a,], Rakshita Pritam Singh Dhar [a], César Pascual García [b,]*, Vihar Georgiev [a,]**

a. University of Glasgow, Glasgow G12 8QQ

b. Luxembourg Institute of Science and Technology (LIST), 41 Rue du Brill, L-4422 Belvaux, Luxembourg

KEYWORDS: Amino Acids, Fingerprints, ISFET, Site-Binding Model, Gouy-Chapman-Stern Model, Surface Potential, Capacitance, Noise, Silanol Sites, Polypeptides

Supplementary Material

A. Behaviour of amphoteric properties of AA transduced by ISFETs.

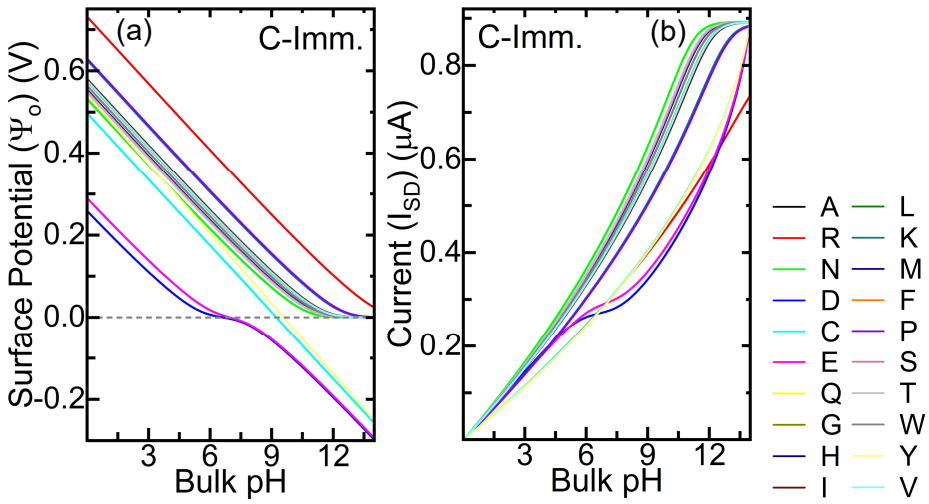


Figure SI 1. (a) Surface Potential (b) Drain Current of ISFET for 20 essential Carboxylic-terminal immobilized amino acids with respect to the pH.

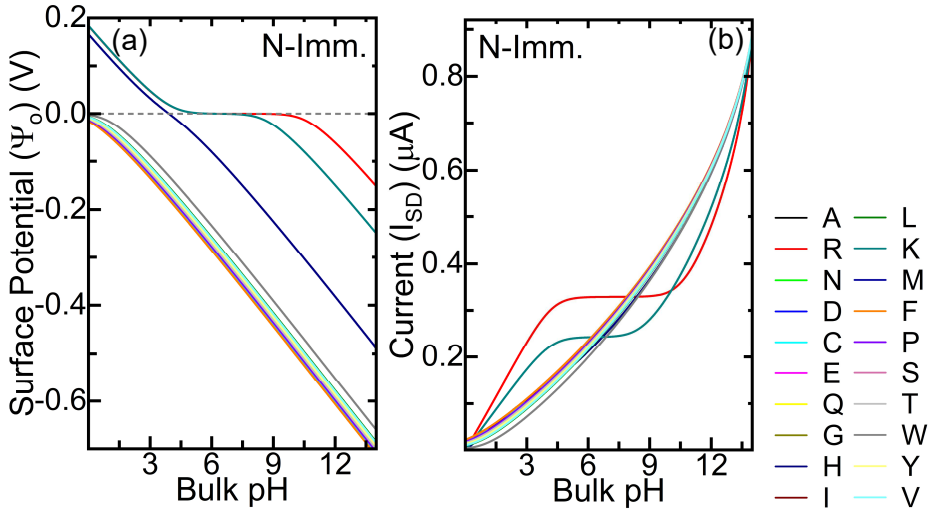


Figure SI 2. (a) Surface Potential (b) Drain Current of ISFET for 20 essential Amine-terminal immobilized amino acids with respect to the pH.

Figures SI 1 and 2 show the surface potential and drain current variation for 20 essential amino acids with respect to the pH immobilized with carboxylic and amine terminal. respectively. Ψ_o was obtained by solving the GCS model and site-binding equation self-consistently which is explained in the methodology section of the manuscript in detail. Throughout the article, including this supporting information, the current was calculated for a large aspect ratio FinFET as explained in the main text. The obtained results show a clear distinction with the C- immobilized amino acids shown in the main text of the article. The proposed approach is flexible enough to account for the superposition of signals due to different types of immobilization over the same surface and helps in deducing the possible number of reactive sites for specific immobilization. It is easier to identify the amino acids due to distinct fingerprints in the surface potential and drain current.

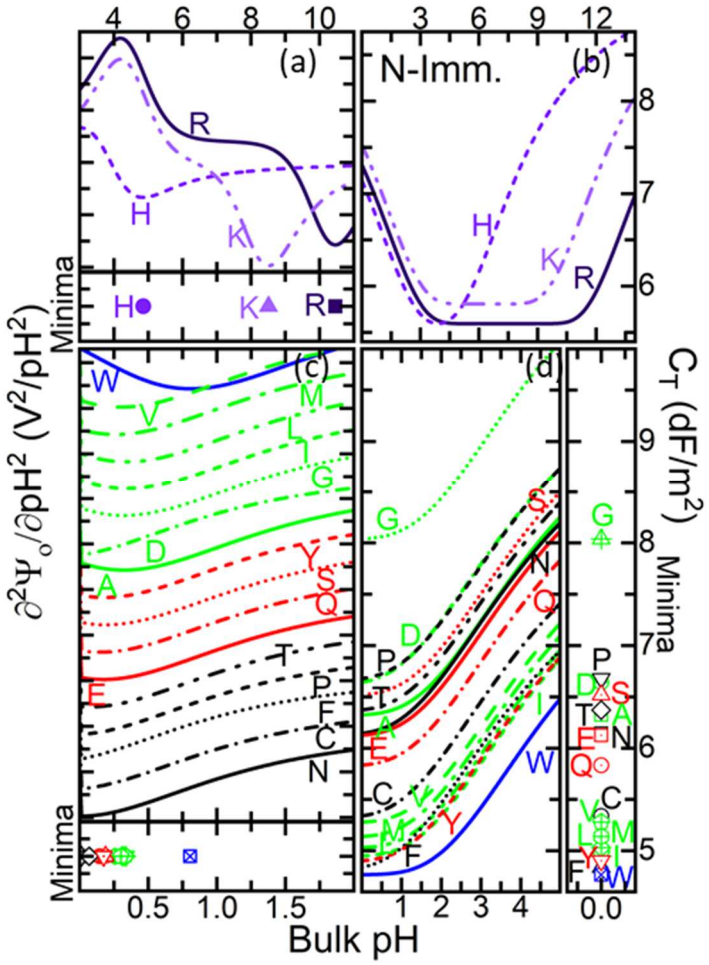


Figure SI 3. (a) and (b) 2nd Gradient of Surface Potential ($\delta\Psi_o^2/\delta pH^2$) and Total Capacitance (C_T) for Carboxyl-immobilized H, K, and R respectively. H, K, and R are shown in blue using a color degradation to identify them in addition to line texture (dashed, double dotted-dashed and solid for H, K and R respectively). (c) and (d) $\delta\Psi_o^2/\delta pH^2$ and C_T for the rest of the Carboxyl-immobilized AAs, respectively. The insets in (a) and (c) show the position of the maxima for each of the AA. The inset in the right of figure (d) show also the minima of C_T. The AAs in (c) and (d) have been grouped in four colors depending on the position of the maxima from lower to higher pH: W is shown in a solid blue line, N, C, F, P and T are shown in black with solid, dashed-dotted, dotted, dashed, and dashed double-dotted respectively. E, Q, S, and Y are shown in Red, solid, dashed-dotted, dotted and dashed respectively. A, D, G, I, L, M and V are shown as green, solid, dashed-dotted, dotted, dashed, double-dotted dashed, long dashed-dotted and long dashed respectively. The notation for the symbols for each AA is the same in (c) and (d).

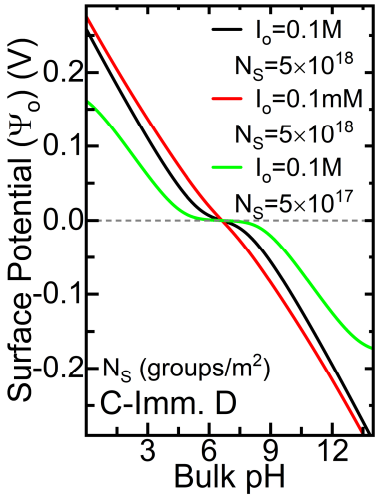


Figure SI 4. Surface Potential of ISFET for different ionic concentrations (I_o) and surface states for Carboxylic-terminal immobilized aspartic acid with respect to the pH.

Figure SI 4 shows the change in the surface potential characteristics with different ionic concentrations and N_S. Even a slight variation in ionic strength, number of immobilized analytes, types of ions, etc, is reflected in the amino acid fingerprints that can help in calibration and benchmarking the possible variables to extract possible information about the vicinity of the analytes.

B. Signal acquisition

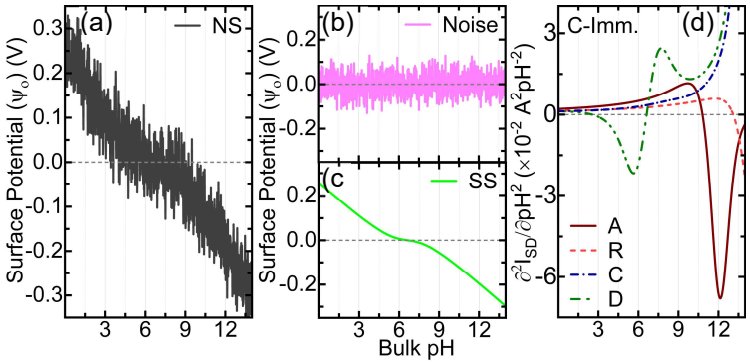


Figure SI 5 (a) Simulation of noisy Ψo vs. pH signal with SNR=10dB; (b) Extracted Noise; (c) Extracted Ψo vs. pH signal from (a) for Carboxy-terminal immobilized aspartic acid; (d) 2nd Gradient of Drain current ($\delta I^2/\delta pH^2$) as a signal response from ISFET for Carboxyl-immobilized AAs (A, R, C, D).

Fingerprint measurements to detect AAs require a high signal-to-noise ratio. Even with precise measurements, any experimentally measured signal of Ψ_o (or output current) may have SNR smaller than 50dB due to intrinsic fluctuations (mainly thermal fluctuations of the ions on the surface and within the electrolyte)[1,2]. Noise affects the recognition of fingerprints as the fast variations perturb the recognition of singularities in the derivates of the signal. Signal processing with noise filtering can be a strategy to improve AA readings and provide access to fingerprints. Traditional filtering techniques such as Moving-mean (Mm), Savitzky-Golay Filtering (SGF), Fast-Fourier Transform (FFT), etc., fail because the smoothing processes may be responsible for the actual data loss due to the dependence on the average number of points, order of the signal, frame width, noise interference with the signal of similar frequency, or the signal is aperiodic.

We propose an alternative for the extraction of actual signals from noisy measurements. To illustrate the process, we introduced a random noise signal with Signal to Noise Ratio (SNR)=10dB (fig. SI 5(a)) simulating

the presence of noise in experiments (fig. SI 5(c)). The noise (fig. SI 5(b)) is extracted by initially subtracting a randomly simulated signal (e.g., using the signal fitted with the model of a wrong AA) from the experiment. In the next step, using FFT, a threshold estimation is extracted by analyzing the power spectral density (PSD). The frequency zones of high amplitude of PSD, corresponding to well-determined frequency are considered signals while the rest of the frequencies with average amplitude are considered as the threshold that will be subtracted. This average noisy FFT components depend also in the SNR of the experimental signal. A detailed process of noise extraction is also shown in supporting information. Figure SI 5(c) shows the extracted signal, which corresponds 100% with the analytical model. In this way, we have shown that it is possible to retrieve the peptide signatures even with SNR < 10dB if the complete titration spectrum of pH can be accessed.

Current output signals can be used to transduce the effects of the Ψ_o and C_T with real advantages to multiplex the signal from many sensors. We calculated the transduction from a FET with a Fin of high-aspect-ratio geometry described in some of our previous works[3,4]. We observed in the calculations that the current varies for different AAs depending on the type of immobilization (by the amino or carboxylic terminals). Here, we show the signal transduction of carboxylic immobilized A, R, C and D in the form of 2nd gradient of the current with respect to the pH shown in fig. SI 5(d). The zero-crossover point of D between the minima and maxima correlates with the isoelectric point of the AA but the same doesn't hold for C due to a consistent increase in current without any significant charge transition for a larger range of pH as the difference in corresponding pK values for protonation and deprotonation (for C) is quite small as compared to D. As for the case of A, the current keeps on increasing till it reaches a saturation point leading to a maximum followed by a minimum of $\delta I^2/\delta pH^2$. A similar characteristic $\delta I^2/\delta pH^2$ is followed by R with a right shift in the curve toward higher pH values due to decreased slope of potential with reduced current per pH sensitivity. Such signal transduction of the AAs is helpful in sensor array multiplexing that can enhance the scope of multi-AAs fingerprint detections while fine-tuning the device characteristics to negate the drift problem[4].

C. Noise filtering methodology for ISFETs.

Here, as an example, we have taken the simulated signal relative to the surface potential as a function of a pH titration for two C-immobilised Aspartic Acid [Figure SI 5(a)- Noisy signal in black]. The used Mm (in green) and SGF (3rd order) (in blue) have 101 samples as the frame width. The threshold for the FFT power spectral density (psd) was $5*10^{-3}$. Figure SI 6(a) shows the Noisy Signal (NS) (black), and a few filtered signals for the C-imm corresponding to the simulations of different noise for 10, 30 and 50 dB for Aspartic Acid. Even with 50dB SNR, the noise present in the filtered signal is bad in terms of accuracy. The noise exists even when the measured signal (with 50dB) is different than C-immobilised Aspartic Acid which makes the whole process of smoothing unreliable when the concern is the accuracy of the measured signal.

We propose an alternative for the extraction of actual signals from the noisy measured signals using the FFT and simulated signals. To show the strength of this approach, we have considered a noisy signal with SNR equal to 10dB. Figure SI 6 (d) shows the Noisy Signal (NS) (black), a randomly simulated Signal (SS) (green)

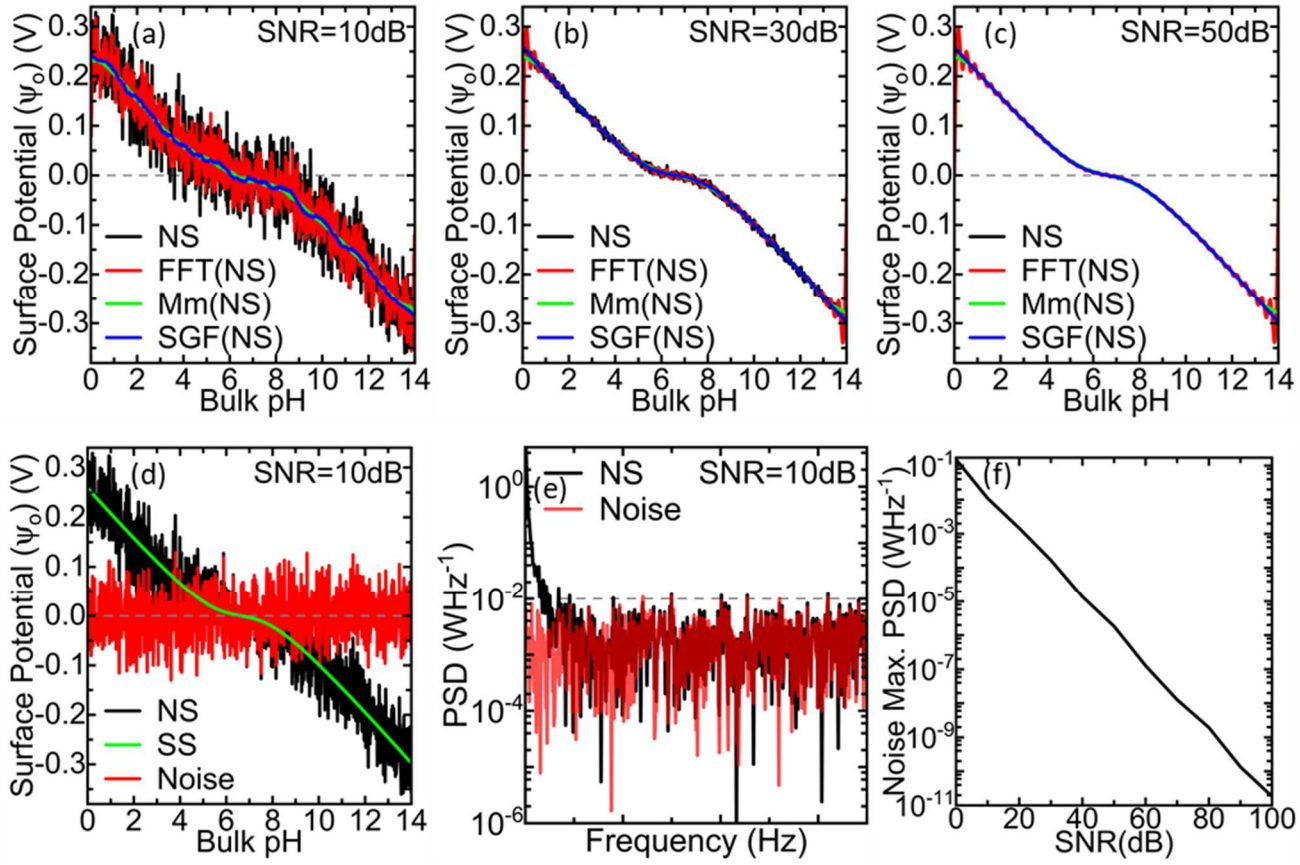
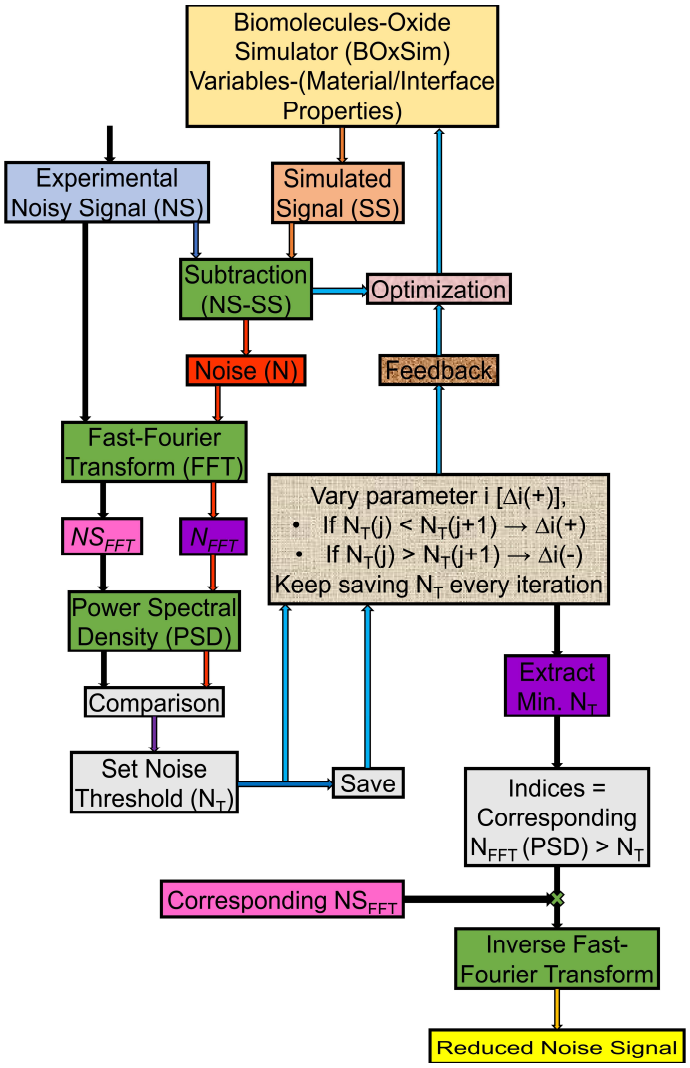


Figure SI 6. Noisy Signal (NS) (black), and Signal filtered using Moving-mean (Mm) (green), Savitzky-Golay Filtering (SGF) (blue) and Fast-Fourier Transform (FFT) (red) for the C-imm. Aspartic Acid for different SNR ratio of (a) 10dB (b) 30dB and (c) 50dB (d) Noisy signal (black) with extracted noise (red) and retrieved original signal (green) (e) Power Spectral Density (PSD) of the Noisy signal (lack) and the extracted noise (red) (f) Max. PSD as the threshold to filter the noise with the variation of SNR

and the extracted noise (red) for C-imm. aspartic acid. The noise is extracted from the noisy signal by subtracting the randomly simulated signal from it. In the next step, using FFT, the power spectral density (psd) of the noise (red) and noisy signal (black) is extracted for the estimation of the threshold (max. psd) selected for the values above 10^{-2} WHz^{-1} shown with the thin dashed reference line in Figure SI 6(e). The remaining signal after the threshold filtering calculated from inverse FFT is then added to the SS for the final output. The max. psd is used as a threshold to filter the noise from the noisy signal to extract the original required signal. The threshold is lower for a measured signal with a higher SNR. Thus, the accuracy of the extracted signal is higher for higher SNR. The threshold (max. psd) varies depending on the measured signal and noise content as shown in Figure SI 6(f). For a 10dB noisy signal, the calculated max. threshold was 10^{-2}. As compared to the previous analysis, the threshold was $5*10^{-3}$ and kept the same for noisy signals with different SNR. Even with a higher threshold, the signal extraction was improper for a 50dB noisy signal as shown in Figure SI 6(f). Thus, the proposed method can be helpful in the accurate calculation of max. psd value for noise filtering. The proposed method can also be used for the proper calibration of the simulated signal. With the included noise, it is hard to depict the comparison between the noisy signal and the simulated signal. Thus, even for an inaccurate simulated signal, the threshold was calculated and compared with different values of threshold calculated from a range of simulated signals depending on the possible variation in input variable while keeping the amplitude within the range of the noisy signal. Thus, using simulated data for the post-processing of the experimental signal can help in providing a highly accurate output.

Block diagram explaining the working of the noise filtering approach



The simulated signal (any random) never interferes with the original signal (noisy signal) but uses a copy of the noisy signal to extract possible noise for the FFT analysis leading to the power spectral density and noise threshold calculations (as per the flow in the block diagram). The noisy signal is used at the end after the extraction of min. noise threshold following the numerous iterations of the random simulated signal compared with the noisy signal. The reduced noise signal is extracted by removing the noise indices from the FFT of the noisy signal. The method is expected to work particularly well to filter the contribution from single species of AA or peptides because they would have a contribution in the same region of the frequency spectrum. However, there is always a possibility of a lost signal while performing the subtraction of the noise [as shown in Figure: a bit of black signal (NS_{FFT}) below the threshold (dash line) and not overlapping the red signal (N_{FFT})] but the lost data account of million times smaller as compared to other filtering techniques. This proof of concept can also be applied for specific real-time noise filtering with the inclusion of a simulator within the experimental measurement system.

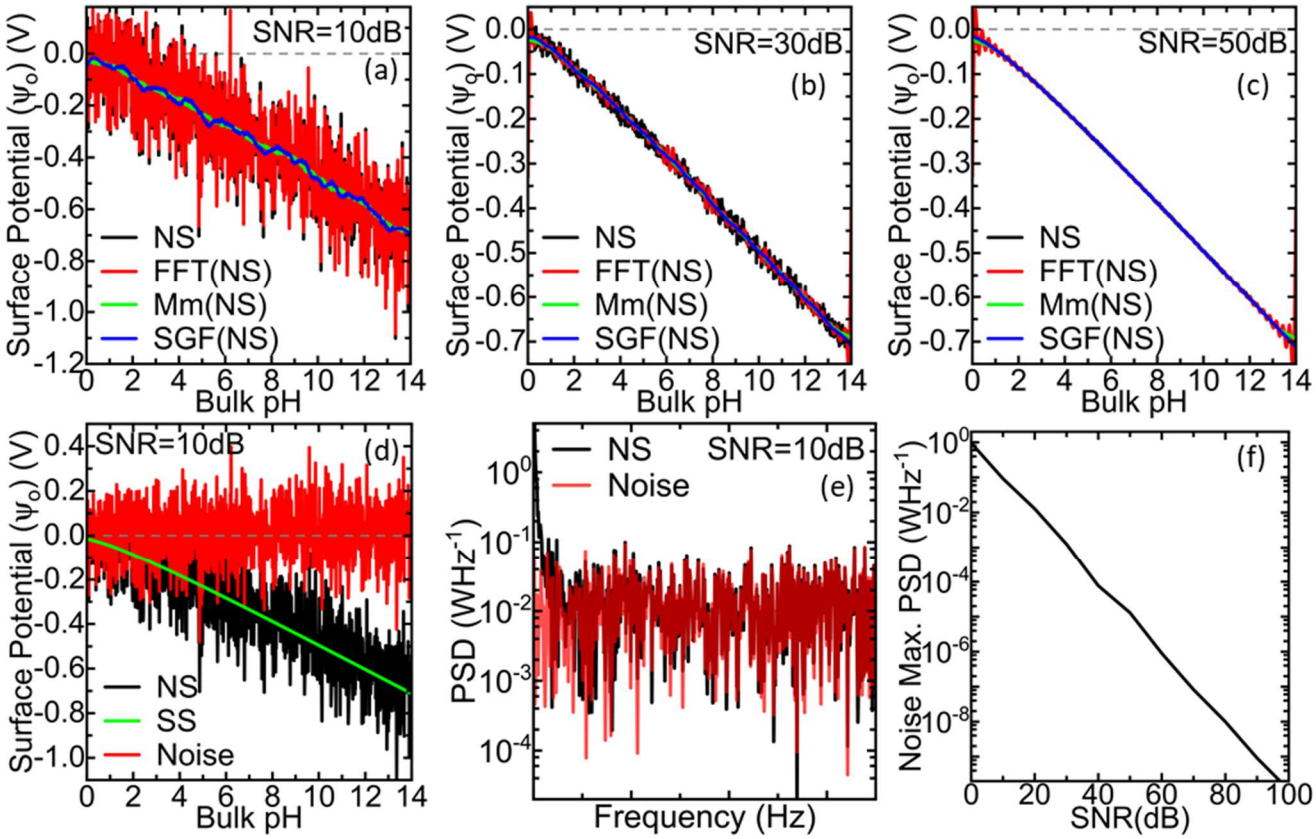


Figure SI 7. Noisy Signal (NS) (black), and Signal filtered using Moving-mean (Mm) (green), Savitzky-Golay Filtering (SGF) (blue) and Fast-Fourier Transform (FFT) (red) for the N-imm Aspartic Acid for different SNR ratio of (a) 10dB (b) 30dB and (c) 50dB (d) Noisy signal (black) with extracted noise and required original signal (d) Power Spectral Density (PSD) of the Noisy signal and the extracted noise (e) Max. PSD as the threshold to filter the noise with the variation of SNR

Figure SI 7 shows the extracted results of the noise filtering approach for amine-terminal immobilized aspartic acid which differs from the C-imm. aspartic acid. The noise threshold changes for every different experimentally measured signal with a variable noise threshold.

D. Chemical interference by the sensor surface

Another relevant experimental condition to reveal the AA fingerprints is the chemical interference from impurities or partial functionalization that may leave proton active radicals not originated on the AAs. For example, semiconductor FETs typically build the sensing interface upon an oxide dielectric in which oxide groups exposed to the electrolyte can also exchange protons, resulting in shifting the isoelectric point. Imperfect passivation of the oxide surface provokes these oxide groups to compete with the proton affinity sites of the AAs, resulting in a chemical interference in the system to study the peptide signatures. To overcome this issue, we can include in the simulation a percentage of non-functionalized groups, which could be calibrated *a priori*. Here, we have studied the case of Carboxyl-terminal immobilized aspartic acid (D) having an amine and carboxylic sidechain with SiO_2 as a gate oxide for the FET. We have simulated different partial functionalizations that result in a percentage of silanol groups from the dielectric barrier interfering with the active reactive sites of the AAs. The total surface states (N_T) are composed of attached analytes (AA) (N_{AA}) and the free silanol sites ($N_{OH} = N_T - N_{AA}$) which are contributing to the surface charge density (σ_o). Figure SI 8(a) shows the simulations for $\delta\Psi_o^2/\delta pH^2$ for Carboxyl-terminal immobilized aspartic acid in the

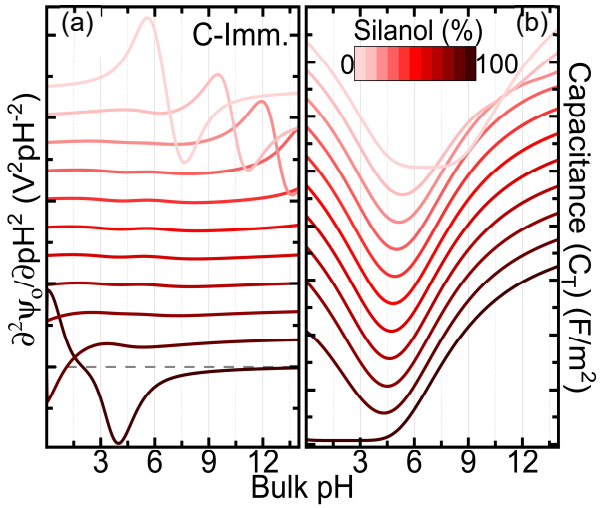


Figure SI 8 (a) and (b) 2nd Gradient of Surface Potential ($\delta\Psi_o^2/\delta pH^2$) and Total Capacitance, respectively, for Carboxyl-immobilized aspartic acid in the presence of silanol sites represented as a percentage of the total surface states. $\delta\Psi_o^2/\delta pH^2$ and CT are plotted with an offset to clarify the minima and maxima for different percentages of silanol.

presence of silanol sites where we represent the percentage of silanol groups with colors degradations from 0 (lighter color, perfect functionalization) to 100% (darker color, pure silica groups) in steps of 10% of the total surface states. The presence of silanol sites changes the surface potential and isoelectric point due to the superposition of charge from the protonated/ deprotonated silanol radicals with the AA depending on the pH value. The silanol site (pK_a = -2) is more reactive to acquire a negative charge as compared to the carboxylic radical of D (pK = 3.65) but silanol site (pK_b = 6) is less effective in protonating as compared to the amine radical of D (pK = 9). As the percentage of the silanol sites increases from a low value (0%) to 10%, the surface potential tends to have lower values due to a more acidic dissociation constant of silanol sites as compared to the carboxyl sites of aspartic acid [Surface potential shown in supplementary material, Figure SI 9]. The zero-crossover point of the double gradient correlates with the isoelectric point of the system with either 100% AA or silanol sites. The isoelectric point shift to acidic values as the percentage of the silanol sites keeps increasing due to the higher effect of smaller dissociation constants of silanol radicals. In the presence of reactive sites from AA and silanol, the zero-crossover point depicts the change in slope of surface potential and the ascendancy of the silanol over the AAs. At lower silanol contributions, D dominates the behaviour in $\delta\Psi_o^2/\delta pH^2$, which are more significant in the range of alkaline pH values as compared to the acidic range due to the early deprotonation of silanol radicals. As the isoelectric point shifts to more acidic values, the maxima of $\delta\Psi_o^2/\delta pH^2$associated with the amino groups however shift to the right to leave space for changes in convexity that occur because of the contributions from the silanol groups to the electrostatic equilibrium with the double layer capacitance. Because of these contributions, the linearity of the surface potential increases flattening $\delta\Psi_o^2/\delta pH^2$, making precisely the most linear contribution at 50% contribution from each of the reactive sites. In the presence of a high contribution of silanol sites that decreases the effective affinity of the surface for protonation as compared to the individual amine sites of D, the behaviour is dominated by the silanol groups and the peaks of $\delta\Psi_o^2/\delta pH^2$ are then observed in the acidic range of pH.

C_T also contributes to the determination of analytes (AAs) by allowing calibration of the contribution of silanol percentage present over the oxide surface. Figure SI 8 (b) shows the variation of total capacitance with respect to pH for C-immobilized aspartic acid in presence of silanol sites. With the increase in silanol percentage, the capacitance decreases at higher pH values and increases at lower pH values with a left shift in the capacitance curve. The capacitance minima correspond to the isoelectric point. The combination of the percentage of silanol groups can thus be determined

by C_T with the position of the isoelectric point and the amplitude variation at the highly acidic/alkaline pH values. The left shift of the capacitance shows the dominance of silanol with an increasing percentage that decreases the effective isoelectric point of the system. The subsequent increase in the flatness of the constant capacitance after 50% (silanol sites) denotes the increased difference in the pK values of the system. This approach confirms the AA's fingerprints with the existence of dominating reactive sites and differentiates from the bare Carboxyl-immobilized AAs with 100% surface coverage.

With the previous calibration of free silanol groups, a correlation between the minima variation of the C_T and the zero-crossover point of the $\delta\Psi_o^2/\delta pH^2$ can help in determining the contribution of silanol groups. For example, if the immobilization of carboxylic groups is done through classical amino-silanes (e.g. (3-Aminopropyl)triethoxysilane), a previous pH titration may be used to determine the number of free silanol groups.

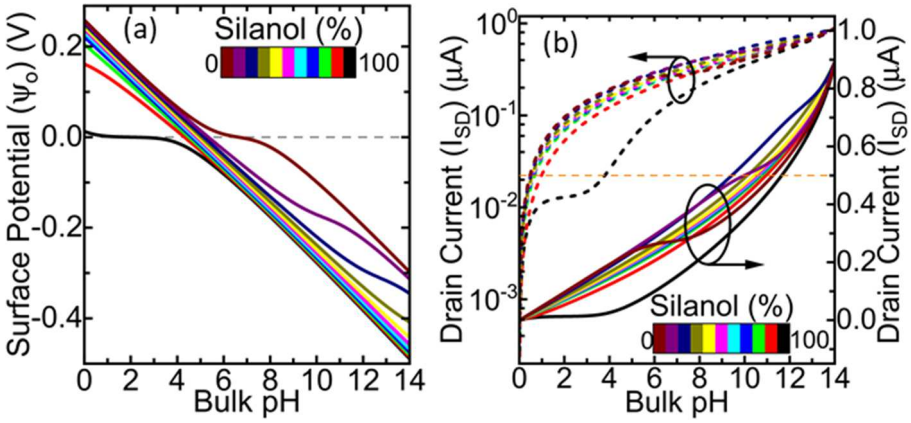


Figure SI 9. (a) Surface Potential and (b) Drain current of ISFET for Carboxylic-terminal immobilized aspartic acid in presence of different percentages of silanol sites with respect to the pH. We show both the logarithm (dotted lines) and linear (solid lines) scales.

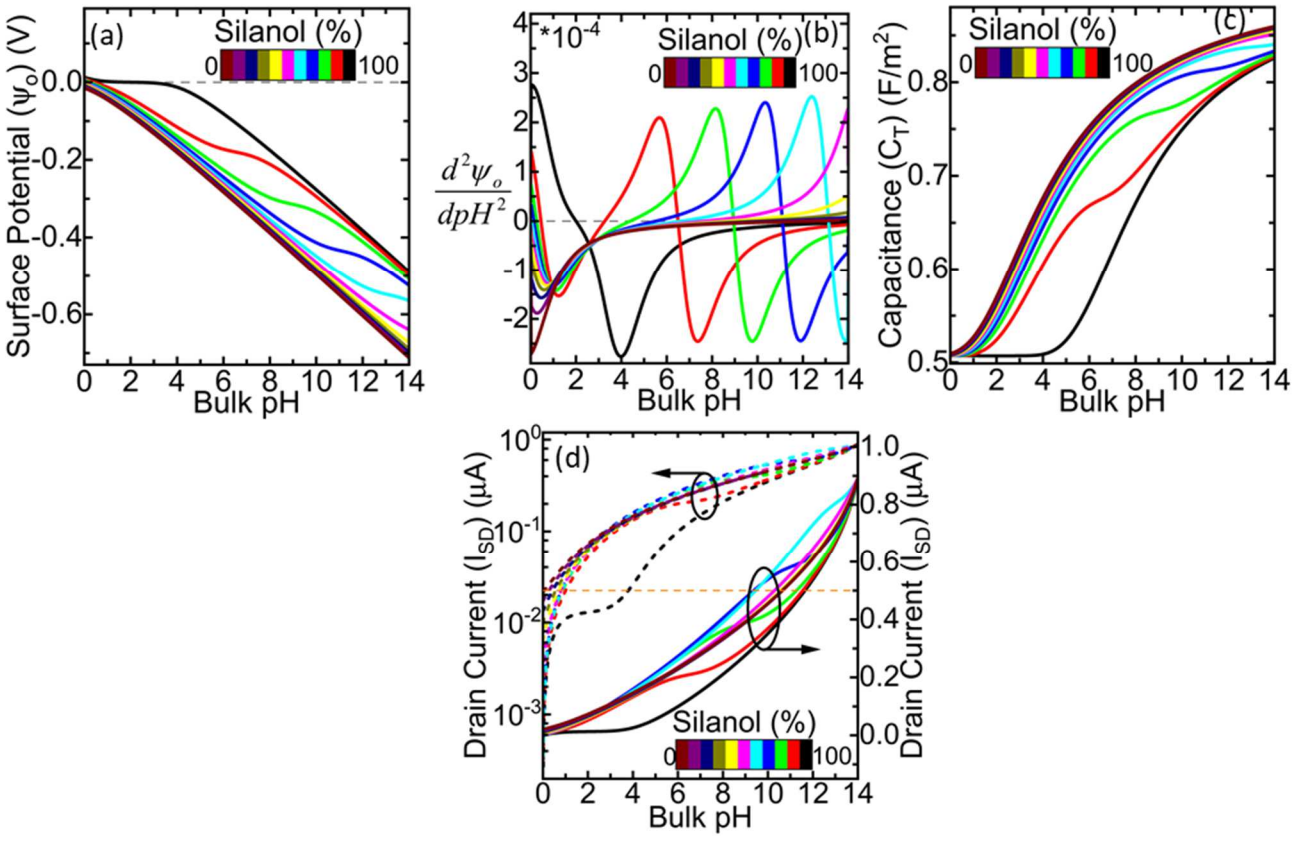


Figure SI 10. (a) Surface Potential (b) 2nd gradient of surface potential (c) Total capacitance (C_T) and (d) Drain current of ISFET for Amine-terminal immobilized aspartic acid in presence of different percentages of the silanol sites with respect to the pH.

Figures SI 9 and SI 10 show the different signals of surface potential (a), second order derivative of the surface potential (b), total capacitance (c) and drain current (d) varying with respect to the pH for Aspartic acid immobilized by carboxylic and amine terminal respectively. The variation in surface potential is different for different immobilization of the aspartic acid in the presence of the same percentage of silanol sites. Thus, depending on the affinity constants of the amino acids or small peptides, the presence of silanol sites will have distinct effects and can be extracted to identify the actual number of amino acids immobilized on the surface.

1. Mele, L. J., Palestri, P. & Selmi, L. General Model and Equivalent Circuit for the Chemical Noise Spectrum Associated to Surface Charge Fluctuation in Potentiometric Sensors. *IEEE Sensors Journal* **21**, 6258–6269 (2021).
2. Go, J., Nair, P. R. & Alam, M. A. Theory of signal and noise in double-gated nanoscale electronic pH sensors. *Journal of Applied Physics* **112**, 034516 (2012).
3. Dhar, R., Kumar, N., Pascual Garcia, C. & Georgiev, V. Assessing the effect of Scaling High-Aspect-Ratio ISFET with Physical Model Interface for Nano-Biosensing Application. *Solid-State Electronics* 108374 (2022) doi:https://doi.org/10.1016/j.sse.2022.108374.
4. Medina-Bailon, C. *et al.* Comprehensive analytical modelling of an absolute pH sensor. *Sensors* **21**, (2021).