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,]**

naveen.kumar@glasgow.ac.uk, r.dhar.1@research.gla.ac.uk, cesar.pascual@list.lu*, vihar.georgiev@glasgow.ac.uk**

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

ABSTRACT: Protein sequencing is key to many biological fields to advance science and prospect medical applications based on proteomics. However, even with advanced techniques such as mass spectrometry it is hard to detect the fluctuations in AA sequences and optical/geometrical isomers, and cannot provide wide access to ex-novo sequencing of low quantities of proteins. We are presenting a new and disruptive method based on FET sensors for AA/polypeptide detection. We show unique signals (fingerprints) of every single AA and polypeptide mutations by solving the site-binding model self-consistently with the Gouy-Chapman-Stern model. The surface potential (Ψ_o), 2nd gradient of the surface potential ($\delta\Psi_o^2/\delta pH^2$) and total surface capacitance (C_T) are used as fingerprint signals to differentiate between the amino acids including the drain current variation as the signal transduction. These fingerprints are based on orthogonal properties of the AAs, which are the different proton affinities of each of the radicals, their dielectric constant and effective length, which is sensitive to the relative positions of the radicals to distinguish between isomers. We studied the generation of signals and proposed a novel noise-filtering technique based on the Fast-Fourier Transform (FFT) and the significantly improved analytical model which is solved

iteratively to reduce data loss. The minimum combined fingerprint resolution is 0.1 units of pH for the separation of singularities found in $\delta\Psi_o^2/\delta pH^2$, linked to the minimum capacitance of the AAs with a needed resolution of 0.01mF/m^2 for surface densities of 10^{14}cm^{-2}, and which can be normalized to lower densities. The effect of noise (>SNR=10dB) and silanol sites can be negated by correlating the AAs signatures from $\delta\Psi_o^2/\delta pH^2$ and capacitance. Thus, the designed methodology and approach can help immensely in designing a new and efficient tool for protein sequencing while solving the problems related to the signal transduction of sensors.

INTRODUCTION

Chemical methods like the Sanger and Edman degradations appeared decades before DNA sequencing[1,2]. However, the supremacy of current Next-Generation DNA Sequencing techniques[3] outperformed proteomics in terms of throughput, costs, and the minimum amount of sample needed for analysis[4] giving genomics the lead of innovation and downgrading ex-novo sequencing to the research of general properties of proteins. Mass spectrometry (MS), the gold standard for protein sequencing, cannot fulfill the requirements to translate to day-to-day applications covering the irrupting biological discoveries that require protein diagnosis[5,6]. In the last decade, the research of new methodologies for protein sequencing has intensified to provide wider access to high throughput analysis of smaller quantities of samples[7–9]. Very significant contributions include combinations of chemical degradations with fluorescence labels[10] and the transduction of selective enzymatic reactions by the detection of light using massively multiplexed metal-oxide-semiconductor circuits[11]. In general, the new alternatives [12,13] are still in the early stages of development facing serious challenges related to the pivotal task of AA identification due to the overlap of the chemical and physical properties of AAs. Although AA identification has

progressed[14–17], new methods to identify the AAs are still a necessity to provide an alternative to MS.

We propose a new way to identify AAs based on the transduction of their amphoteric properties with field-effect transistors (FETs) that can generalize access to protein sequencing. Using FETs to distinguish proteins by their isoelectric point and studying the variations of the surface potential was already considered by the inventor of the ISFET, P.Bergveld[11], but it was discarded because the protein charges often lay above the Debye length and thus are screened by the ions in solution. However, AAs have a shorter length in comparison to proteins. The new materials used in the current FET fabrication[18,19] and techniques of protein manipulation[7,9,19,20] provide prospects for the analyses of low quantities of samples, providing a way to transduce the AA fingerprints.

In this paper, we describe a new simulation methodology and analytical results, which are disruptive because in addition to provide enough resolution to distinguish different AAs, our new method is also able to read the signal from small polypeptides to discover mutations and provide information on the sequence. This could avoid having to degrade the whole peptide and provide a methodology similar to shotgun MS. Using a site-binding theory for the electrochemical response of FETs, we discuss for the first time, the signatures that appear in the surface potential or surface capacitance from a single amino acid immobilized on the surface of FET sensors. We discuss the experimental conditions required to resolve the electrical signal in terms of electrical and chemical noise. We suggest new fabrication strategies that bring AA recognition to a new level with the current state-of-the-art technologies (e.g., pure 2D materials[21] and FETs with high-K dielectrics[22]).

METHODOLOGY

Our simulation methodology is based on a self-consistent method provided by the Site-Binding equations and Gouy-Chapman-Stern model[23] used to describe two measurable quantities as the surface capacitance (C_T) and surface potential (Ψ_o), which depend on orthogonal properties such as the charge, the dielectric constant and the length of the AAs. The charges at the solid-liquid interface are considered from the contributions of immobilized AAs or polypeptides using the site-binding theory. We have assumed a surface coverage (N_S) for single species of AA or small peptides immobilized at the interface between the dielectric barrier[24] of the FET and the electrolyte[25]. Similar to reactive sites on an oxide, each AA or small peptide has different radicals that contribute to the total charge of the molecule depending on the external pH value[26]. The positive and negative surface charge density contributions (σ_+ and σ_-, respectively) for single AA or a polypeptide can be calculated from the contributions of proton affinities or dissociation constants of each radical (k_i);

$$\sigma_+ = qN_S\left(\frac{\sum_{i=0}^{(n-1)}\left([H^+]_S^{(n+m)-i} * k_i * k_{i-1} * \dots\dots\dots * k_{i-(n+m-1)} * k_{i-(n+m)}\right)}{\sum_{i=0}^{(n+m)}\left([H^+]_S^{(n+m)-i} * k_i * k_{i-1} * \dots\dots\dots * k_{i-(n+m-1)} * k_{i-(n+m)}\right)}\right) \quad (1)$$

$$\sigma_- = qN_S\left(\frac{-\sum_{i=0}^{(m-1)}\left([H^+]_S^{i} * k_{(n+m)-i} * k_{(n+m)-(i+1)} * \dots\dots * k_0\right)}{\sum_{i=0}^{(n+m)}\left([H^+]_S^{(n+m)-i} * k_i * k_{i-1} * \dots\dots\dots * k_{i-(n+m-1)} * k_{i-(n+m)}\right)}\right) \quad (2)$$

$$where, k_1 < k_2 < k_3 < k_{n+m}, k_0 = 1, k_{j<0} = 1$$

where $[H^+]_S$ is the surface proton concentration, which depends on the bulk proton concentration ($[H^+]_B = 10^{-pH_B}$, where pH_B is the bulk pH value) and the surface potential (Ψ_o) with a Boltzmann relation $\left([H^+]_S = [H^+]_B e^{\left(-\frac{q\Psi_o}{kT}\right)}\right)$. In our model, we can consider AAs immobilized from the Carboxylic (C-) or Amine (N-) terminals by cancelling the corresponding affinity (or dissociation) constants. For this work, we have used the standard pK values in the literature[27,28] of

the molecules in a liquid. The interaction with different sensor substrates can modify these pK values, which could be later considered phenomenologically to refine the model[29].

The charge on the liquid interface surface is balanced with the charge in the diffuse layer and the charge in the semiconductor channel. The GCS model describes the electrical double layer (EDL) formed on the oxide surface consisting of an oxide-electrolyte interface, stern layer and the diffuse layer. The stern layer represents the effective uncharged (dielectric) region between the surface charge and counter-ions. The averaged permittivity of the stern layer varies with the charge density in its vicinity, ionic content of the electrolyte and several other factors but for simplicity of the analytical model, the stern capacitance (C_{stern} = 0.8F/m^2) is kept constant as used in Van Hal, 1996[23]. However, even with variable C_{stern} that may lead to the change in potential amplitude due to series connection with diffuse layer capacitance, the fingerprints that we have discovered will still be present with slightly different values that can be observed by the phenomenological correction compensated by the effective proton affinities. The surface potential within the GCS model is the summation of the potential drop across the stern layer (V_{stern}) and the potential at the shear plane (zeta potential, Ψ_ζ):

$$\Psi_o = V_{stern} + \Psi_\zeta = \frac{\sigma_{DL}}{C_{stern}} + \Psi_\zeta \qquad (3)$$

$$\sigma_{DL} = Q_o sinh\left(\frac{q\Psi_\zeta}{2k_BT}\right), \qquad where\ Q_o = \sqrt{8k_BT\varepsilon_w I_o N_{avo}} \qquad (4)$$

Here, q, k_B, T, ε_w, I_o and N_{avo} represent the electronic charge, Boltzmann constant, temperature, permittivity of electrolyte, ionic concentration, and Avogadro's number respectively. Whereas C_{stern} and σ_{DL} are the stern capacitance and diffuse-layer charge density respectively. σ_{DL} in the

GCS model is equated to $\sigma_o = \sigma_+ + \sigma_-$ described by the site-binding model to obtain the values of Ψ_ζ and Ψ_o, which are solved self-consistently.

Along with charge, the length (l_{eff}) and dielectric permittivity (ε_{eff}) of the AAs also contribute in generating a unique signature in terms of intrinsic capacitance ($C_{int}=\varepsilon_{eff}/l_{eff}$). The total capacitance (C_T) is calculated from the series combination of intrinsic capacitance (C_{int}), diffuse layer capacitance (C_{DL}) and C_{stern}:

$$\frac{1}{C_T} = \frac{1}{C_{int}} + \frac{1}{C_{DL}} + \frac{1}{C_{stern}}, \qquad where\ C_{DL} = \frac{\partial \sigma_{DL}}{\partial \Psi_\zeta} \tag{5}$$

By using the total surface, we calculated the surface potential and the capacitance from the Gouy-Chapman Stern model[23] for full coverage (using $5\text{x}10^{14}$ groups/cm^2). The superposition of surface potential with the reference gate bias (V_G) controls the FET characteristics by changing the depletion width (W_D) in the semiconductor channel.

$$\Psi_o + V_G = V_{OX} + V_S, \qquad where \quad V_{OX} = \frac{qN_A W_D}{C_{OX}}\ and\ V_S = \frac{qN_A W_D^2}{2\varepsilon_{Si}} \tag{6}$$

$$I_{SD} = \frac{A}{\rho L} V_{SD}, \qquad where \quad A = h(W - W_D) \tag{7}$$

where, N_A, A, L, h, ε_{Si} and W are the channel doping, cross-sectional area, channel length, height of the channel, permittivity of the silicon and device width respectively. Whereas V_{OX}, V_S, C_{OX}, I_{SD}, and V_{SD} are the potential drop across the gate oxide, potential drop across the semiconductor channel, gate oxide capacitance, source-to-drain current and source-drain bias respectively. Finally, we used the continuity equation of surface potential that equals the potential drop across the oxide and semiconductor with a reduced value of gate bias to calculate the current response in sensors with a FinFET geometry[30] studied in our recent publication[31,32].

RESULTS

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

To pursue a conceptually different method of identification of AAs, we considered the physicochemical properties of proteins. AAs have different isoelectric properties, swapping charges as a function of pH in a unique way. Following the acid titration's pH of a solution of AAs can provide their fingerprints, which is a classical textbook experiment[33–35]. However, the identification using this method requires an equivalent number of AA molecules to the titrated acid, which makes it impractical for the typical samples where the concentration of proteins is usually low. Our approach to transduce the amphoteric properties of AA is to use FETs because they can be sensitive to charge changes even with very few AAs approaching single-molecule using graphene-FET[36–38]. Therefore, here, we will calculate the surface potential and capacitance properties of AAs and small peptides as a function of pH. Our approach does not require the calibration of the absolute surface potential of the FET sensor because it relies on the detection or the proton affinity (pK) of each radical of the AA which produces singularity points on the surface potential when pH is modulated. In our model, we consider a covalent immobilization of AAs by one of the Carboxyl- or Amino-terminals that eliminates the affinity point corresponding to the new bond with the solid phase surface[39,40]. To provide a proof of concept of our approach and to demonstrate the AA fingerprints, we can first disregard the modifications of proton affinities that occur with the interaction of the substrate, which later can be determined phenomenologically.

We illustrate the signatures with four exemplary AAs: Alanine (Ala or A), Arginine (Arg or R), Cysteine (Cys or C) and Aspartic acid (Asp or D). As represented by the molecular structures depicted in fig.1, AAs can exhibit charged/neutral states depending on the acidity: at pH values

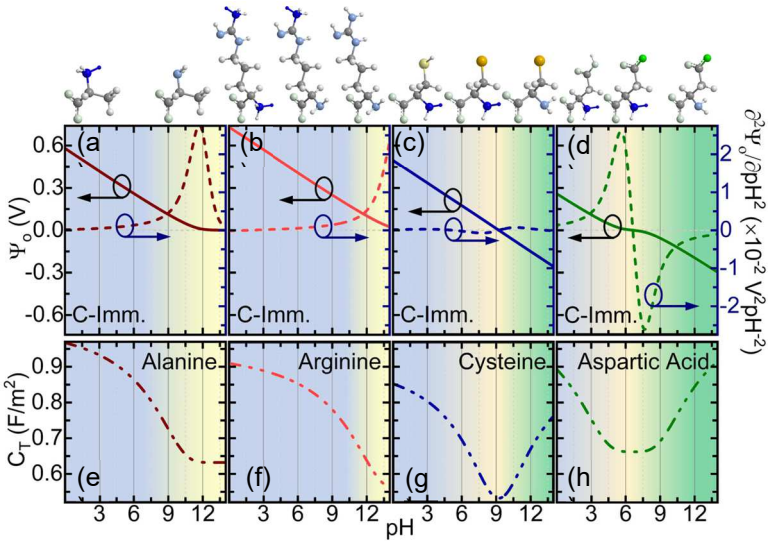

Figure 1 (a) to (d) Surface Potential (solid line) and 2nd Gradient of Surface Potential (δΨo2/δpH2) (dashed line) for A, R, C and D amino acids, respectively (e) to (h) Total Capacitance of ISFET with respect to the pH for Carboxy-terminal immobilized amino acids for A, R, C and D, respectively.

higher than the pK of the amine-terminal the amino groups tend to lose their extra hydrogen to become neutral, while at pH values lower than the pK, the group tends to remain protonated. Meanwhile, at more acidic pH than their pK values, the carboxylic/thiol-sidechains of AAs remain neutral, and at a more alkaline solution, these groups become negatively charged due to the deprotonation. The AAs in fig.1 are chosen to demonstrate the possibility of different charges due to various functional groups (including sidechains). Navy blue color in fig.1 represents positively charged amine radicals, green and gold colors represent the negative carboxylic and thiol radicals, respectively. We have represented the uncharged radicals with the corresponding lighter colors. Figures 1(a) to (d) show the behavior of Ψ_o with the change in pH of the electrolyte for the Carboxy-terminal immobilized A, R, C and D, respectively (shown with brown, red, blue and green solid lines, respectively). A is a bifunctional AA with a free amine radical (pK = 9.69) which tends to remain protonated even at high pH values leading to a positive Ψ_o (the values of Ψ_o are calculated with respect to the bulk electrolyte) in the full pH range. Protonation and deprotonation are continuous processes (subtle changes), which reflect fluctuations averaged over several molecules and time. Statically, some amine radicals deprotonate as pH increases towards the pK

value resulting in decreasing Ψ_o approaching asymptotically to zero. Although the complete deprotonation of A is never reached due to the action of the double layer capacitance and the Bolzman distribution of protons. However, at very alkaline values the charge decreases considerably approaching the isoelectric point, which is observed in the loss of sensitivity. R is a trifunctional AA with two free amine radicals (pK = 9.04, 12.48) whose protonated states are responsible for higher charging values as compared to A. Even with early deprotonation of the amine radical connected to the α-Carbon of R, the protonated amine radical connected to the sidechain keeps the positive charge leading to a higher Ψ_o even at pH=14, with linear behaviour in the full range. D is a trifunctional AA that after the immobilization by the C- terminal remains with a free amine radical (pK = 9.60) and a sidechain of the carboxylic group (pK = 3.65). D acquires a positive charge at lower PH values due to the protonated amine and neutral carboxylic group. With an increase in pH, Ψ_o decreases from positive to negative values due to the deprotonation of the amine and carboxylic radicals changing the overall charge of the molecule and allowing the AA to have an isoelectric point in the zwitterion state, which happens at the pH value which is the average of the pK values of amine and carboxylic radicals. The most prominent feature of Ψ_o is the isoelectric, which shows as the least sensitive part with a zero in the first and second-order derivatives of Ψ_o. C is also a trifunctional AA which after C-immobilisation has a thiol sidechain (pK = 8.18) and amine radical (pK = 10.28). The thiol group acquires negative charges similar to carboxylic radicals but at a higher pH value. The Ψ_o for C (blue solid line fig.1(c)) varies similarly to D but with more linear characteristics due to the smaller difference in dissociation constants of free radicals, which didn't allow the charge stabilization during the transition..

The results in fig.1 were calculated for an ionic strength of 0.1M. Lower ionic strengths show more linear behaviour of the surface potential with respect to pH, thus a less defined shape of the isoelectric point. The better resolution of the isoelectric point at higher ionic strengths is a consequence of the screening effect that compensates for the changes in the double layer capacitance, allowing the Ψ_o to be closer to the charging states defined by the pK values. Away from the isoelectric point, Ψ_o is changing almost linearly with the bulk pH due to the contribution of the double-layer capacitance which acts to equilibrate the charges at the sensor/electrolyte interface. The slope of Ψ_o in the linear regime is associated with the density of active radicals. Following the identification of bifunctional and trifunctional amino acids, fitting Ψ_o in its linear region to the site-binding model provides the number of active sites (N_s), which can later be used to normalize the capacitance. The calculations in fig.1 considered $N_s = 5 \times 10^{14}$ groups/cm^2. Lower N_s result in a more extreme shape of the isoelectric point (flatter potential and thus less sensitivity in the isoelectric point) and also the observation of a saturation effect at extreme pHs. Both, the effect of the ionic strength and the increasing number of protons, are shown in the supplementary material [Figure SI 4].

A better resolution of the AA fingerprints on Ψ_o is also found in the second derivative of the surface potential ($\delta\Psi_o^2/\delta pH^2$) shown in dashed lines in fig. 1(a) to (d). The convex to concave transition in the curve Ψ_o vs. pH associated with the isoelectric point showed as a zero in $\delta\Psi_o^2/\delta pH^2$ signaling, representing the point of minimum sensitivity, which is observed for C and D (fig.1(c) and (d), respectively). $\delta\Psi_o^2/\delta pH^2$ also displays other singularity points that appear as maxima or minima associated with the protonation/deprotonation of the radicals appearing between the isoelectric point and the pH values corresponding to the dissociation constants of the radicals of each AA, but that is affected by the screening of ionic charges and the Boltzmann distribution of the

concentration of protons. For the case of D, a deprotonated sidechain of carboxylic radical coexists with the protonated amine in the range of 3.65<pH<6.675 with more number of protonated amine radicals controlling the charge transition near the isoelectric point. The effect of C's deprotonated thiol radical compared to the amine radical in D is more subtle. The variations of proton affinities fall in the range 8.18<pH<9.23 leading to a bit of sudden charge transition resulting in minima before and maxima after the isoelectric point as the opposite of D. The singularities associated with the protonation and deprotonation (like the maxima and minima of D and C) are shifted from the exact pK values by the double-layer capacitance and the effect of the interplay between the buffering activity of the AA, which shifts the local pH. A higher value of the sidechain pK shifts the $\delta\Psi_o^2/\delta pH^2$ curve to higher pH values for AAs with similar behavior, whereas the maxima and minima values are higher for a larger difference in amine-pK and sidechain-pK. Figure 1(a) shows the absence of an isoelectric point, but the deprotonation of A's amine radical reaches 50% at its pK value and decreases drastically with a further increase in pH due to the higher charge balance effect of counterions with reduced buffer capacity. As compared to A, an extra amine sidechain of R adds up the positive charge in superposition to the amine radical (α-Carbon) resulting in a shift of the maxima to a higher PH. These singularity points at the zero and the maxima and minima shown by $\delta\Psi_o^2/\delta pH^2$ can constitute the variable to construct unique fingerprints of each AAs.

To improve the unique fingerprints from each AA, we also considered the dependence of C_T on pH. While Ψ_o depends mainly on the distribution of charges between the double layer capacitance and the AA themselves, C_T depends also on the dielectric constant and thickness of the AAs, which are orthogonal properties to the charge. C_T can be derived with electrical impedance spectroscopy from the FET transconductance at variable frequencies. The effective electrolyte-gate capacitance is comprised of the different contributions in series: (i) a device capacitance, (ii) AA/peptide

capacitance, and (iii) the double layer capacitance. As the double-layer capacitance also depends on the different pK of the AAs, C_T also exhibits unique features that will depend on the pH. Due to the elimination of the double layer capacitance, the isoelectric point will be observed as a minimum in the C_T vs pH. The values of those minima can be normalized to the N_S derived from the linear part of Ψ_o. The normalized C_T minima are associated with the length and the dielectric constant of each neutral peptide, establishing the second orthogonal signature (fingerprint) of each AA, which can be considered together with the pH behavior as the C_T vs pH or its derivative that exhibit similar singularity to the case of Ψ_o.

Figures 1(e) to (h) display C_T vs. pH for the same four AAs. Being the shortest AA, A shows higher capacitance across the pH range as compared to R (longest). C shows the smallest capacitance at the isoelectric point due to the contribution of sulfur which has a smaller dielectric permittivity, whereas a higher intrinsic capacitance of D increases C_T near the isoelectric point and extreme pH values as the shape of the capacitance curve depends on the C_{DL}. The constant capacitance range is larger for AAs with a bigger difference in carboxy and amine pK values due to the higher probability of charge reduction with balanced protonation and deprotonation of reactive sites. It is worth noting the difference in the expression of the isoelectric point in C_T for C and D compared to its expression in Ψ_o. While C exhibits a very linear behavior and the isoelectric point is less apparent in Ψ_o, the shape of the minimum is well defined in C_T. Instead, for D, while the isoelectric point in Ψ_o is more apparent, C_T shows a broader less resolved minimum. The results reveal that the higher away capacitances from the isoelectric point for D and C are due to the charge stored in the electrical double layer but the capacitance for R and A keeps on decreasing with an increase in pH due to saturation of deprotonated amine-reactive sites. Considering the normalization to N_S in the range of 10^{10} groups/cm^2 (accessible in experimental conditions[42]), the

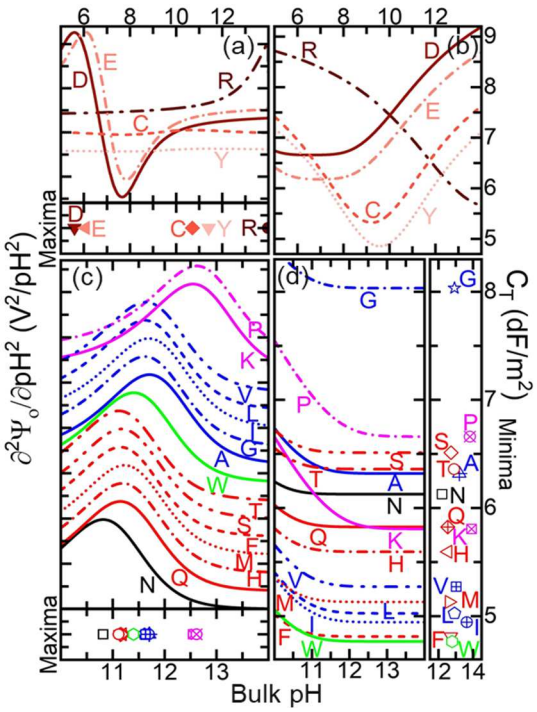

Figure 2(a) and (b) 2nd Gradient of Surface Potential (δΨo2/δpH2) and Total Capacitance (CT) for Carboxyl-immobilized R, D, C, E, and Y respectively. R, D, C, E, and Y are shown in brown using a color degradation to identify them in addition to line texture (dotted-dashed, solid, short dashed, dotted short-dashed and dotted for R, D, C, E, and Y, respectively). (c) and (d) δΨo2/δpH2 and CT 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 fig (d) show also the minima of CT. The AAs in (c) and (d) have been grouped in six colors depending on the position of the maxima from lower to higher pH: N is shown in a solid black line, Q, H, M, F, S and T are shown in red with solid, dashed-dotted, dotted, dashed, dashed double dotted, and long dashed-dotted, respectively. W is shown in a solid green line A, G, I, L and V are shown as blue, solid, dashed-dotted, dotted, dashed, and double-dotted dashed, respectively. K and P are shown in pink with solid and dashed-dotted lines, respectively. The notation for the symbols for each AA is the same in (c) and (d).

normalized capacitance requires a resolution of 10 μF to identify the AAs, which is a well achievable measurable quantity. Considering a FET gate of 1 μm^2, also easily attainable for current microfabrication techniques, this signal would correspond to an average of 100 molecules, which is largely improving the state of the art for MS (~10,000 molecules needed for ex-novo sequencing). A higher-order capacitance gradient with respect to pH can pinpoint the pK values as zero-crossover points with stored charge transition from positive to negative. We can use our approach to iteratively solve the surface potential with self-consistency convergence for precise calculation of the active surface states (N_S)/dissociation constants and have a full characterization of the peptide interphase which can be compared with the experiments.

B. Amino Acid fingerprints

We consider now the distinction of the fingerprints from the 20 essential AAs immobilized by Carboxyl- terminals shown in fig.2 (the fingerprints of the AAs immobilized by the amino-terminals are also shown in SI). The analysis is divided into two categories: one considering the AAs with sidechain resulting in separated and distinct maxima/minima of $\delta\Psi_o^2/\delta pH^2$ (fig.2(a)) with their corresponding minima in the C_T (fig.2(b)) and the other group considering the AAs without the sidechains which each exhibiting a single maximum of $\delta\Psi_o^2/\delta pH^2$ between pH 10 and 14 (fig.2(c)) and their corresponding C_T (fig.2(d)). The isoelectric points of the first group (fig. 2(a)) are always displayed as zeros on $\delta\Psi_o^2/\delta pH^2$ which characterizes the AAs after the immobilization that has zwitterionic behavior: D, E, C and Y for the Carboxyl-immobilised AAs. Carboxyl-immobilised R can also be added to this list of readily identified AAs due to the extra amino group that doubles the sensitivity and thus, provides a distinct behavior to the surface potential. In addition to the isoelectric point, the existence of both maxima and minima of D, E, C and Y are well defined between pH 5 and 12 (shown in the inset of fig.1(a)), which univocally identifies these AAs. In addition, the position of the maximum of R above pH 14 also clearly identifies this AA. Also, the position of the minima of C_T is shown in fig. 2(b) can be used to distinguish D, E, C and Y by their isoelectric point, and R by the extreme alkaline minimum that it exhibits.

Figure 2(c) shows the maxima of $\delta\Psi_o^2/\delta pH^2$ for the AAs considering Carboxyl-immobilisation, which lay between $10.5 < pH < 13$. These AAs have only a single amino group active, which does not allow them to have a real isoelectric point within a pH range 0-14. The maximum of $\delta\Psi_o^2/\delta pH^2$ is associated with the dominating radical, which provides a distinct way to characterize some of the AAs. The $\delta\Psi_o^2/\delta pH^2$ maxima associated with N and W (shown in black and green, respectively) are at pH 10.8 and 11.4 (fig. 2(c)) and can be distinguished from the rest by at least 0.3 pH units.

K and P (shown in pink with solid and dotted dashed lines) are also well separated from the other groups at pH 12.6. The C_T can be used to distinguish amongst them as their minima lay at 0.58 and 0.67 F/m^2, respectively (shown in the inset of fig. 2(d) with a square with diagonal cross and dotted circle signs, respectively). There are two groups of AAs with very similar behavior of the $\delta\Psi_o^2/\delta pH^2$ vs. pH amongst them, with a group of AAs with a maximum in pH ~11.2 plotted in red (Q, H, M, F, S and T plotted with solid, dashed-dotted, dotted, dashed, dashed double dotted, and long dashed-dotted lines, respectively) and 11.7 plotted in blue (A, G, I, L and V plotted with solid, dashed-dotted, dotted, dashed, and double-dotted dashed, respectively). Although the exact pH of the maxima/minima is different in all AAs, a clearer distinction of the AAs is provided by C_T, which is enough to distinguish most of these AAs with standard experimental conditions assuming the normalization of C_T to N_S. fig. 2(d) shows the capacitance variation, and the later inset shows the capacitance minima. Hence, it can be concluded that capacitance curves can be used in combination with the surface potential data or independently to characterize each AA. However, a few of the intracategory AAs have a lower capacitance resolution of 0.01mF/m^2 with similar C_T minima for Leucine and Isoleucine, due to their difference in length that occurs because of the different relative positions of the carboxylic and amino groups. This calculation is an approximation since we do not consider the changes in pK values that occur after immobilization, but assuming the value, which could be changed after a phenomenological adjustment of the values, thus the change in capacitance would allow distinguishing two isomers that cannot be resolved by current MS methods. Our model also neglects the 3D effect of the position of the charges, which would increase the differences between molecules, in particular considering that each radical would have differences in the local screening.

C. Signal acquisition

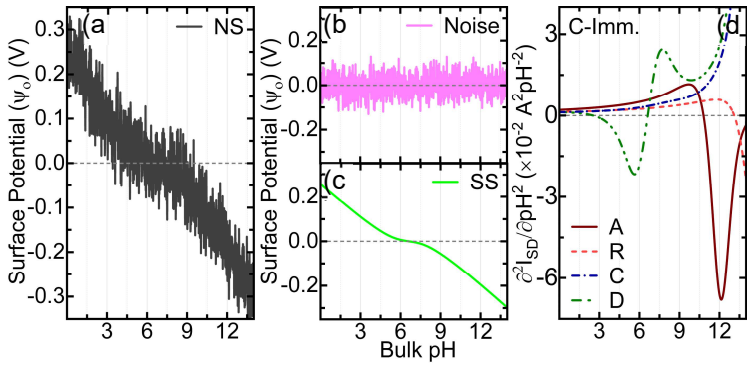

Figure 3 (a) Simulation of noisy Ψo vs. pH signal with SNR=10dB; (b) Extracted Noise; (c) Extracted Ψo vs. pH signal from (a) for Carboxy-terminal immobilized aspartic acid; (d) 2nd Gradient of Drain current (δI2/δpH2) as a signal response from ISFET for Carboxyl-immobilized AAs (A, R, C, D).

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)[43,44]. 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. 3(a)) simulating the presence of noise in experiments (fig. 3(c)). The noise (fig. 3(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 3(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[31,32]. 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. 3(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[31].

D. Chemical interference by the sensor surface

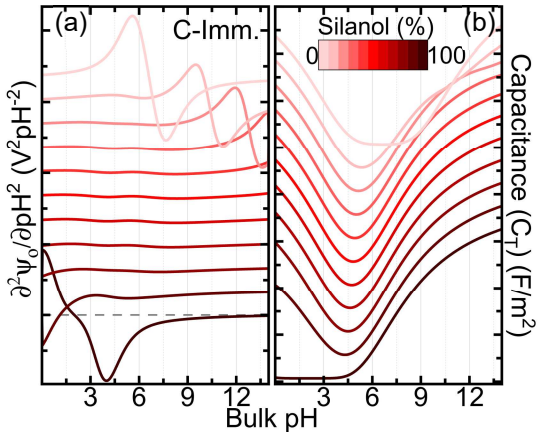

Figure 4 (a) and (b) 2nd Gradient of Surface Potential (δΨo2/δpH2) and Total Capacitance, respectively, for Carboxyl-immobilized aspartic acid in the presence of silanol sites represented as a percentage of the total surface states. δΨo2/δpH2 and CT are plotted with an offset to clarify the minima and maxima for different percentages of silanol.

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 4(a) shows the simulations for $\delta\Psi_o^2/\delta pH^2$ for Carboxyl-terminal immobilized aspartic acid in the 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 7]. 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 4(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.

E. Signatures of small peptides

For some applications, the discovery of signal changes within similar peptide sequences may be a more efficient way of characterization rather than complete ex-Novo sequencing[45]. The addition

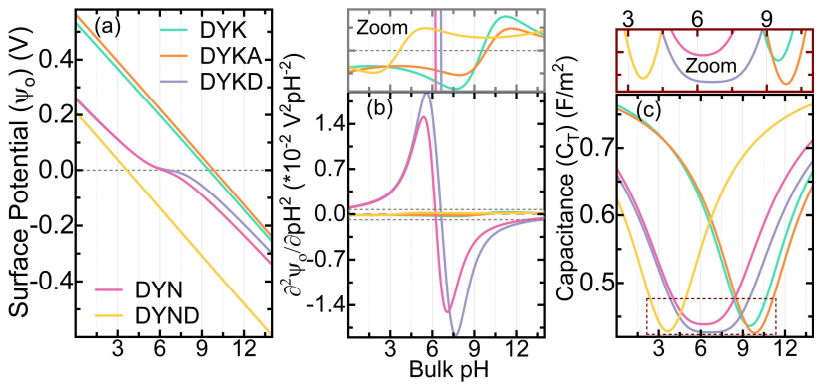

Figure 5 Left-hand side: (a) Surface Potential; bottom middle: (b) 2nd Gradient of Surface Potential (δΨo2/δpH2); top middle: (c) Zoomed δΨo2/δpH2 of the selected area (grey color); bottom left-hand side: (d) Total Capacitance; top left-hand side: (e) Zoomed CT of the selected area (wine color) of ISFET with respect to the pH for Amino-terminal immobilized polypeptides (DYK) and derivatives

of AAs to a peptide increases the complexity of the charge distribution as the electro-affinity of residues and terminals are modified by each bond making difficult to predict the surface potential or the total capacitance. The problem of modeling the final behavior is indeed complex as it involves other interactions including hydrogen, van-der-Waals and electrostatic forces, which are all highly correlated. Here, we propose a phenomenological way to distinguish variations in short polypeptides (addition, substitution or mutation) which is based on the detection of charges present on the reactive sites of the polypeptides using the site-binding method as discussed in the methodology section. We propose here to study the DYK sequence because it can be recognized by flag monoclonal antibodies[46,47]. We consider the peptide immobilized by the carboxylic-terminal of the first AA, and thus only the amine-terminal of the last AA and the sidechains of the AAs contribute to different proton affinities. The values of pK in this system could also be resolved phenomenologically. To discover new signatures, we use the same approach until now of considering the nominal pK values of the AAs in liquid. However, the phenomenological constants can be expected to provide more resolution than the nominal ones, as the complex interactions would affect the phenomenological pK values including the 3D distribution of the charges that are affected by different screenings of the double layer capacitance. The proposed model uses the individual affinity of reactive sites present on different AAs of a polypeptide chain with the

effective length and average permittivity to calculate the surface potential as the sensing signal, and the values have been considered from the original AAs without modifications of the bonds as a first-order approximation. The effective length of the polypeptides is calculated using Chemdraw® for stable 3D structures with minimum energy between the two extreme ends. The actual length (perpendicular to the surface) of the polypeptide may be even smaller due to immobilization with C-terminal and orientation. Here, we have used 3/4 AAs [D-Aspartic Acid, Y-Tyrosine, K-Lysine, A-Alanine, N-Asparagine] based polypeptides to check for signatures of peptide modifications. As the length of the peptide chain is roughly half of the Debye screening length for the ionic strength in our calculations, we still consider that our model can provide a qualitative approach that can predict fingerprints of Ψ_o and C_T expected for the variations here considered. The peptides are chosen to differentiate between a growing peptide, mutations and substitutions in a peptide. Out of the used samples, DYN and DYND would not be recognized by the flag monoclonal antibody.

Figure 5(a) shows the surface potential varying with respect to the pH for DYK and four other co-related polypeptides. DYK (immobilized by its C-terminal) results in a rich peptide as holds three side chains in addition to the N-terminal because the three AAs are trifunctional. The immobilization by D results in two carboxylic groups (the side chains from D and Y) and two amino groups from K. The Ψ_o of DYK (shown in cyan in fig.5(a)) appears almost linear due to the rich contribution from amino and carboxylic groups. The isoelectric point is around pH 10 due to the strong proton affinity of amine groups and the alkaline dissociation constant of Tyrosine. When observed in detail (zoom region shown in the inset of fig.5(b)), $\delta\Psi_o^2/\delta pH^2$ shows the zero crossing of the isoelectric point and a rich behaviour with a local maximum and minimum at more acidic values where the inflexion points are between the isoelectric point and the values of the radicals

with more acidic pK, and another maximum between the isoelectric point and the most alkaline pK value corresponding to the influence of the N-terminal of the Lysine compensated with the action of the double layer capacitance. The addition of A in the modified peptide DYKA (shown in orange) shows similar behaviour. Ψ_o shows slightly higher values and a small shift of the isoelectric point towards more alkaline values (visible in the inset of fig. 5(b)) due to a higher proton affinity of the amine-reactive site of A as compared to K. This small difference between DYKA is expected as A is the simplest bifunctional AA, which only increases slightly the length of the peptide, which can be detected in the decrease of the values of C_T, that also reflects the small shift of the isoelectric point (fig. 5(c)).

Instead, the addition of D (instead of A) to conform DYKD (shown in blue) modifies the equilibrium of radicals by adding an extra carboxylic group with lower pK than the one of the Tyrosine. As a result of the more negative charges, Ψ_o decreases with respect to DYKA or DYK (fig. 5(a)). Moreover, the extra side chain makes the isoelectric point more apparent because of the increasing difference between the proton affinities of the carboxylic and the amino groups. This behaviour results in the $\delta\Psi_o^2/\delta pH^2$ shown in fig.5(b) almost similar to aspartic acid because of the acidic split of the radicals. However, the differentiation with the single AA is clear due to the decrease in C_T, as shown in fig.5(c).

A similar amphoteric effect to the addition in the DYKD is obtained with the substitution of N in the DYK to conform DNK. Similarly, to the action of the addition of D in the DYKD sequence, the substitution of K by N equilibrates the proton affinity of the radicals decreasing the Ψ_o of DYN (shown in pink in fig.5) because the proton affinity of the amino sidechain of Asparagine does not have as high proton affinity due to the close presence of a carboxyl group (fig.5(a)). The amphoteric equilibrium results in the centered isoelectric point and a $\delta\Psi_o^2/\delta pH^2$ behaviour of DYN

similar to DYKD (fig. 5(b)). The distinction between DYKD and DYN can be measured by C_T, in this case, changes mostly due to the length of the AA resulting in higher capacitance for the shorter peptide DYN with respect to DYKD (fig. 5(c)). The addition of aspartic acid to DYN to conform DYND (shown in yellow) again breaks the equilibria between carboxylic and amino groups with the addition of an extra carboxylic group, which confers this sequence to the most negative charging behavior of all of these AAs, shown by the lowest surface potential and the most acidic isoelectric point (fig. 5(a)). Again the different pK values dispersed in all the acidic ranges, provide this molecule with high linearity of Ψ_o but rich $\delta\Psi_o^2/\delta pH^2$ (shown in fig. 5(b)), which is opposite to the sequences DYK and DYKA showing only a minimum in the pH range more acidic than the isoelectric point, and a maximum and a minimum at pH more alkaline than the isoelectric point. Also, C_T in fig. 5(c) falls in the range of the other tetra-amino acid peptides.

All in all, the behaviour of Ψ_o, $\delta\Psi_o^2/\delta pH^2$, and C_T from these peptides can be explained based on the properties of the proton affinity of their radicals and their length (as we have avoided AA with sulphur), and although they have a richer behaviour than single AAs that may require more complex analysis, still, their fingerprints can be retrieved from their analysis. As the length of the peptides is below the screening length, we expect to be able to measure these peptides' characteristics. Nevertheless, optimizing the ionic strength in these analyses may be crucial because while decreasing it may have a positive effect to decrease the Debye length to probe longer peptides, it can also have a negative impact on the manifestation of the fingerprints by increasing the linearity of the Ψ_o observed. Measuring polypeptides instead of single AAs can help in multiplexing several FET sensors to detect possible variations in a polypeptide or protein-peptide interactions even with different samples. Thus, irrespective of the complexity of the polypeptides, using a reaction like Edman's degradation to remove N- peptides, with a detailed analysis of

surface potential and capacitance including the signal transduction from FET can bring out the precise tool for peptide sequencing.

DISCUSSION OF THE RESULTS

In this paper, we have proposed a method to discover the signal of single AA which is a significant leap in respect to the other Mass Spectrometry alternatives because it is based on simple FET technology that can transduce the fingerprints from orthogonal properties such as the amphoteric behaviour, the dielectric constant and the length of AAs. We have derived an analytical model using the site-binding method modified to facilitate the recognition of AAs or small polypeptides or to distinguish changes within the sequence, in the form of surface potential, capacitance and current, provided that the target analyte remains within the Debye length. The minima/maxima of the $\delta\Psi_o^2/\delta pH^2$ and capacitance clarify the signatures of the different amino acids, whereas the pH value corresponds to the zero-crossover of the $\delta\Psi_o^2/\delta pH^2$ and minima of the capacitance matches the isoelectric point from pH-titration.

Further possibilities open as the pH and other external parameters like the temperature and the ionic strength also modify the interaction forces that result in the properties probed by our method. Having different strategies to immobilize peptides by their alpha carboxylic or amine terminals would also contribute to creating a multi-dimensional space of surface potential and capacitance with the temperature and acidity parameters where each peptide would have a complex signature measured by FETs. The correlation of $\delta\Psi_o^2/\delta pH^2$ and capacitance can help in the extraction of the AA fingerprints and the number of surface states even in presence of silanol sites. The derived methodology of using the analytical model in a combination of FFT solved iteratively with

parametric variations can smoothen the noisy signal with the possibility of the data loss less than a million times smaller than other standardized noise filtering techniques.

In summary, we have designed a simulation framework that can help in the advancement of single AA detection or protein sequencing using ISFETs as a cost-effective and efficient alternative. The proposed approach allows the benchmarking of the amino acid or polypeptide fingerprints irrespective of the drift and noise generated within the system-under-test.

We believe this work proposes a ground-breaking approach to protein sequencing since the amino-acid recognition depends on three orthogonal properties (the proton affinity, the dielectric constant and the length of the AAs) different in all the AAs that can be measurable with any FET with two combined measurable quantities (Ψ_o and C_T). In particular, it is worth mentioning the perspectives of two kinds of FETs: high-K dielectric semiconductor FETs and graphene FETs. The use of High-K dielectrics in the interface gives better chances to measure because it improves the capacitance coupling of Ψ_o with the output current, improving the fidelity of to the surface potential and decreasing the noise. Furthermore, graphene biosensors can be advantageous because of the direct coupling of the surface potential with the Dirac point bringing the capacitance coupling to the quantum limit. The quantum coupling of the surface potential with the Dirac points improves also the sensitivity to resolve very few molecules approaching a single molecule, almost independently of the size of the sensor. Also, graphene is a material insensitive to pH[48], which avoids the interference described in our article. Novel approaches and the revival of classical techniques of solid phase degradation of peptides together with the advance of microfluidic techniques to manipulate minimal quantities of peptides, open perspectives to improve the current state-of-the-art of protein sequencing.

CONFLICTS OF INTEREST

The authors declare they have no conflicts of interest to publish this investigation.

FUNDING

This project was funded by the European Innovation Council under the Horizon 2020 program of Future emerging technologies. Project: 862539 — ElectroMed.

REFERENCES

1. Niall, H. D. [36] Automated edman degradation: The protein sequenator. in *Methods in Enzymology* vol. 27 942–1010 (Academic Press, 1973).
2. Heather, J. M. & Chain, B. The sequence of sequencers: The history of sequencing DNA. *Genomics* **107**, 1–8 (2016).
3. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nature Biotechnology* **26**, 1135–1145 (2008).
4. Aslam, B., Basit, M., Nisar, M. A., Khurshid, M. & Rasool, M. H. Proteomics: Technologies and their applications. *Journal of Chromatographic Science* vol. 55 182–196 (2017).
5. Winston, T. & Gregory, T. Beyond mass spectrometry, the next step in proteomics. *Science Advances* **6**, eaax8978 (2022).
6. Bell, A. W. *et al.* A HUPO test sample study reveals common problems in mass spectrometry–based proteomics. *Nature Methods* **6**, 423–430 (2009).
7. Swaminathan, J. *et al.* Highly parallel single-molecule identification of proteins in zeptomole-scale mixtures. *Nature Biotechnology* **36**, 1076–1082 (2018).
8. Restrepo-Pérez, L., Joo, C. & Dekker, C. Paving the way to single-molecule protein sequencing. *Nature Nanotechnology* **13**, 786–796 (2018).
9. Alfaro, J. A. *et al.* The emerging landscape of single-molecule protein sequencing technologies. *Nature Methods* **18**, 604–617 (2021).
10. Schneider, A. F. L. & Hackenberger, C. P. R. Fluorescent labelling in living cells. *Current Opinion in Biotechnology* **48**, 61–68 (2017).
11. Bergveld, P., van Hal, R. & Eijkel, J. *The remarkable similarity between the acid-base properties of ISFETs and proteins and the consequences for the design of ISFET biosensors*. *Biosensors & Bioelectronics* vol. 10 (1995).
12. Pavlopoulou, A. & Michalopoulos, I. State-of-the-art bioinformatics protein structure prediction tools (Review). *International Journal of Molecular Medicine* vol. 28 295–310 (2011).
13. Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-resolution microscopy with DNA-PAINT. *Nature Protocols* **12**, 1198–1228 (2017).

14. di Muccio, G., Rossini, A. E., di Marino, D., Zollo, G. & Chinappi, M. Insights into protein sequencing with an α-Hemolysin nanopore by atomistic simulations. *Scientific Reports* **9**, (2019).

15. Civitarese, T. & Zollo, G. Triggering Amino Acid Detection by Atomistic Resolved Tunneling Current Signals in Graphene Nanoribbon Devices for Peptide Sequencing. *ACS Applied Nano Materials* **4**, 363–371 (2021).

16. Kennedy, E., Dong, Z., Tennant, C. & Timp, G. Reading the primary structure of a protein with 0.07 nm3 resolution using a subnanometre-diameter pore. *Nature Nanotechnology* **11**, 968–976 (2016).

17. Ohshiro, T. *et al.* Detection of post-translational modifications in single peptides using electron tunnelling currents. *Nature Nanotechnology* **9**, 835–840 (2014).

18. Dhar, R. P. S., Kumar, N., Medina-Bailon, C., García, C. P. & Georigiev, V. P. TCAD Simulations of High-Aspect-Ratio Nano-biosensor for Label-Free Sensing Application. in *2021 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EuroSOI-ULIS)* 1–4 (2021). doi:10.1109/EuroSOI-ULIS53016.2021.9560701.

19. Reed, B. D. *et al.* Real-time dynamic single-molecule protein sequencing on an integrated semiconductor device. *bioRxiv* 2022.01.04.475002 (2022) doi:10.1101/2022.01.04.475002.

20. Hong, J. M. *et al.* ProtSeq: Toward high-throughput, single-molecule protein sequencing via amino acid conversion into DNA barcodes. *iScience* **25**, 103586 (2022).

21. Zhang, P. *et al.* Modulation the electronic property of 2D monolayer MoS2 by amino acid. *Applied Materials Today* **14**, 151–158 (2019).

22. Leshanskaya, L. I. *et al.* Towards understanding the origin of the hysteresis effects and threshold voltage shift in organic field-effect transistors based on the electrochemically grown AlOx dielectric. *Thin Solid Films* **649**, 7–11 (2018).

23. van Hal, R. E. G., Eijkel, J. C. T. & Bergveld, P. A general model to describe the electrostatic potential at electrolyte oxide interfaces. *Advances in Colloid and Interface Science* **69**, 31–62 (1996).

24. Sciuto, E. L. *et al.* Functionalization of bulk SiO2 surface with biomolecules for sensing applications: Structural and functional characterizations. *Chemosensors* vol. 6 (2018).

25. Besar, K., Ardoña, H. A. M., Tovar, J. D. & Katz, H. E. Demonstration of Hole Transport and Voltage Equilibration in Self-Assembled π-Conjugated Peptide Nanostructures Using Field-Effect Transistor Architectures. *ACS Nano* **9**, 12401–12409 (2015).

26. Beroza, P., Fredkin, D. R., Okamura, M. Y. & Feher, G. Electrostatic calculations of amino acid titration and electron transfer, Q-AQB-->QAQ-B, in the reaction center. *Biophys J* **68**, 2233–2250 (1995).

27. *CRC Handbook of Chemistry and Physics*. (CRC Press, 2016). doi:10.1201/9781315380476.

28. Pahari, S., Sun, L. & Alexov, E. PKAD: a database of experimentally measured pKa values of ionizable groups in proteins. *Database* **2019**, baz024 (2019).

29. Lorieau, J. L., Louis, J. M. & Bax, A. Helical Hairpin Structure of Influenza Hemagglutinin Fusion Peptide Stabilized by Charge−Dipole Interactions between the N-Terminal Amino Group and the Second Helix. *J Am Chem Soc* **133**, 2824–2827 (2011).

30. Rani, D., Rollo, S., Olthuis, W., Krishnamoorthy, S. & García, C. P. Combining chemical functionalization and finfet geometry for field effect sensors as accessible technology to optimize pH sensing. *Chemosensors* **9**, 1–12 (2021).

31. Medina-Bailon, C. *et al.* Comprehensive analytical modelling of an absolute pH sensor. *Sensors* **21**, (2021).

32. 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.

33. Xiao, Xu & Tao. Conductance Titration of Single-Peptide Molecules. *J Am Chem Soc* **126**, 5370–5371 (2004).

34. Taylor, W. H. *Formol Titration: An Evaluation of its Various Modif icaiions*. *AN EVALUATION OF* vol. 82.

35. 'Harris, L. J. '. The titration of amino- and carboxyl-groups in amino-acids, polypeptides, etc. Parts I—III.—Investigations with aqueous solutions. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character* **95**, 440–484 (1923).

36. Puster, M. *et al.* Cross-Talk Between Ionic and Nanoribbon Current Signals in Graphene Nanoribbon-Nanopore Sensors for Single-Molecule Detection. *Small* **11**, 6309–6316 (2015).

37. Guan, J. *et al.* Direct single-molecule dynamic detection of chemical reactions. *Science Advances* **4**, (2018).

38. Arjmandi-Tash, H., Belyaeva, L. A. & Schneider, G. F. Single molecule detection with graphene and other two-dimensional materials: nanopores and beyond. *Chemical Society Reviews* **45**, 476–493 (2016).

39. Tullman, J., Callahan, N., Ellington, B., Kelman, Z. & Marino, J. P. Engineering ClpS for selective and enhanced N-terminal amino acid binding. *Applied Microbiology and Biotechnology* **103**, 2621–2633 (2019).

40. Raliski, B. K., Howard, C. A. & Young, D. D. Site-Specific Protein Immobilization Using Unnatural Amino Acids. *Bioconjugate Chemistry* **25**, 1916–1920 (2014).

41. Salgın, S., Salgın, U. & Bahadır, S. *Zeta Potentials and Isoelectric Points of Biomolecules: The Effects of Ion Types and Ionic Strengths*. *Int. J. Electrochem. Sci* vol. 7 www.electrochemsci.org (2012).

42. Militano, F. *et al.* Influence of protein bulk properties on membrane surface coverage during immobilization. *Colloids and Surfaces B: Biointerfaces* **143**, 309–317 (2016).

43. 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).

44. 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).

45. Vyatkina, K. *et al.* De Novo Sequencing of Peptides from High-Resolution Bottom-Up Tandem Mass Spectra using Top-Down Intended Methods. *PROTEOMICS* **17**, 1600321 (2017).

46. Schmidt, P. M. *et al.* Taking down the FLAG! How Insect Cell Expression Challenges an Established Tag-System. *PLOS ONE* **7**, e37779- (2012).

47. Srila, W. & Yamabhai, M. Identification of Amino Acid Residues Responsible for the Binding to Anti-FLAG^{TM} M2 Antibody Using a Phage Display Combinatorial Peptide Library. *Applied Biochemistry and Biotechnology* **171**, 583–589 (2013).

48. Fu, W. *et al.* Graphene Transistors Are Insensitive to pH Changes in Solution. *Nano Letters* **11**, 3597–3600 (2011).