A Journal of the Gesellschaft Deutscher Chemiker

Angewandte International Edition GDCh Chemie

www.angewandte.org

Accepted Article

Title: Multiplex discrimination of single amino acid residues in polypeptides by single SERS hot spot

Authors: Jian-An Huang, Mansoureh Z. Mousavi, Giorgia Giovannini, Yingqi Zhao, Aliaksandr Hubarevich, Miguel A. Soler, Walter Rocchia, Denis Garoli, and Francesco De Angelis

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.202000489

Link to VoR: https://doi.org/10.1002/anie.202000489

WILEY-VCH

10.1002/anie.202000489

WILEY-VCH

RESEARCH ARTICLE

Multiplex discrimination of single amino acid residues in polypeptides by single SERS hot spot

Jian-An Huang [a]*, Mansoureh Z. Mousavi [a], Giorgia Giovannini [a], ¶, Yingqi Zhao [a], Aliaksandr Hubarevich [a], Miguel A. Soler [b], Walter Rocchia [b], Denis Garoli [a], [c]*, Francesco De Angelis [a]

[a] Dr. J.A. Huang, Dr. M.Z. Mousavi, Dr. G. Giovannini, Dr. Y.Q. Zhao, Dr. A. Hubarevich, Dr. D. Garoli, Dr. F. De Angelis
Plasmon Nanotechology Unit,
Istituto Italiano di Tecnologia,
Via Morego 30, 16163 Genova, Italy
E-mail: Jian-an.huang@iit.it, Denis.garoli@iit.it

[b] Dr. M.A. Soler, Dr. W. Rocchia
CONCEPT Lab,
Istituto Italiano di Tecnologia,
Via Melen 83, 16152 Genova, Italy

[c] Dr. D. Garoli
AB ANALITICA s.r.l., Via Svizzera 16, 35127 Padova, Italy

¶ Dr. G. Giovannini,
Present address: EMPA Federal Swiss research Institute, St. Gallen, 9014, Switzerland

Supporting information for this article is given via a link at the end of the document

Abstract: Surface-enhanced Raman spectroscopy (SERS) is a sensitive label-free optical method that shows potential for single-molecule sequencing. However, SERS detection of protein sequences with single-residue sensitivity suffers from signal dominance of aromatic amino acid residues and backbones, impeding detection of non-aromatic amino acid residues. Here, we trap a gold nanoparticle in a plasmonic nanohole to generate a single SERS hot spot for single-molecule detection of 2 alike polypeptides (vasopressin and oxytocin) and 10 distinct amino acids that constitute the 2 polypeptides. Significantly, both aromatic and non-aromatic amino acids are detected and discriminated at single-molecule level either at individual amino acid molecules or within the polypeptide chains. Correlated with Molecule Dynamics simulation, our results suggest that the signal dominance due to large spatial occupancy of aromatic rings of the polypeptide sidechains on gold surfaces can be overcome by the high localization of the single hot spot. The superior spectral and spatial discriminative power of our approach can be applied to single-protein analysis, fingerprinting and sequencing.

Introduction

The advent of analytical techniques with extremely low limits of detection has led to dramatic progress in the field of single-molecule sensing and in particular in nucleic acids sequencing[1]. In particular, revolutionary advances in DNA sequencing technology based on nanopores have been achieved during the last decades[2], but the research interest on this topic is still enormous. The next technological revolution would be a fast and reliable analysis and sequencing of proteins[3] that are the primary actors in virtually all life processes. Coded by DNA sequences known as genes, proteins can yield far more compelling information than may be gleaned from DNA alone. Single-molecule analysis of proteins would be of enormous value by offering the potential to detect, for example, low-abundance proteins that are overshadowed by high-abundance proteins in mass spectroscopic measurements of biological fluids[4].

One major limitation in single-molecule sequencing of proteins is the discrimination of the 20 distinct amino acids that are the building blocks of proteins. The electrical readout method that was successful to discern 4 distinct current levels for nucleobases in single DNA faces a big challenge to recognize the 20 distinct current signals corresponding to the amino acids[3]. Alternatively, integration of optical detection methods such as fluorescence[5] and surface-enhanced Raman spectroscopy (SERS)[6, 7] have shown interesting potential in complementing the electrical readouts. Significantly, plasmonic nanopores are demonstrated to offer improvements to not only detection but also slowing down protein translocation.[8] In a recent work, we demonstrated that a plasmonic nanopore allows multiplex SERS discrimination of single DNA with single-base sensitivity[7].

SERS relies on plasmonic metal nanostructures that exhibit an amplified and localized electromagnetic field on their surface (so-called hot spot) upon illumination of their resonant wavelength. When analyte molecules enter a hot spot, the latter excites and enhances the Raman signals of the molecules to emit SERS spectra. Even though a hot spot usually covers more than one molecule (multi-molecule), the SERS spectra can distinguish their species as well as count their amount in the hot spot due to their narrow fingerprint Raman peaks.[9] Such capability of multiplex discrimination has been demonstrated to detect single biomolecules such as DNA and proteins[10, 11].

A key step to single-protein sequencing by SERS is discrimination of single amino acid residues in a protein chain. Nevertheless, SERS spectra of proteins and amino acids suffer from two problems. The first one is that the SERS spectra of proteins are

dominated by the spectra of aromatic amino acid residues and backbones[11-15]. Non-aromatic amino acids are invisible in the protein spectra. Among the 20 types of amino acids, only 4 of them have aromatic rings. As a result, plenty of structural information of the proteins is missing unless single-molecule spectra of non-aromatic amino acid residues in a protein are identified. Moreover, SERS detection of individual non-aromatic amino acids at single-molecule level are still missing, probably due to their small Raman cross-sections and fast movement in solution. Secondly, SERS spectra of both proteins and individual amino acids exhibit peaks much broader than those in normal Raman spectra, hindering the multiplex discrimination of amino acids in the protein chains.[13, 14] The peak broadening is partially due to SERS signals collected from many hot spots of the plasmonic nanostructures on which the analyte molecules are adsorbed with different conformations.[14] Therefore, a SERS system able to detect single non-aromatic amino acids from a single hot spot is desired to lay a foundation towards single-molecule protein analysis and sequencing.

Recently, we developed an electro-plasmonic system able to generate a single strong SERS hot spot to discriminate single DNA bases in oligonucleotides[7]. Here, we use the same electro-plasmonic system to explore the residue difference at the single-molecule level between two alike and important polypeptides, Vasopressin (VAS) and Oxytocin (OXT), which consist of nine amino acids and differ by two distinct amino acid residues[16]. Briefly, analyte molecules are firstly adsorbed on gold nanostars (AuNSs), which are then trapped in a gold nanohole to create a single SERS hot spot for exciting the molecules and producing single-molecule SERS spectra. We show that the approach is able to discriminate VAS from OXT by the Raman spectra of the amino acid residues, including the non-aromatic ones. To further demonstrate the capabilities of the method, we collect single-molecule SERS spectra of 10 types of amino acids that are the building blocks of the polypeptides. Both aromatic and non-aromatic amino acids are successfully detected with small peak widths as results from single amino acids excited by the intense hot spot. This result proves that Raman Spectroscopy enhanced by a plasmonic hot spot can provide most of the features necessary to achieve protein detection or sequencing at single-molecule level without the trouble of labelling. These features are namely: 1) it can reach spatial resolution comparable to single amino acids; 2) it can reach single-amino-acid sensitivity; 3) it has the spectroscopic discrimination power necessary to determine which amino acid is flowing through the hot spot. Finally, the spectra are analyzed with Molecule Dynamics simulation to evaluate residue conformations of the polypeptides adsorbed on the AuNSs, which may shed light on the aforementioned two problems of protein SERS.

Results and Discussion

Single-molecule spectra of individual amino acids by single hot spot.

The electro-plasmonic system consists of a gold nanohole array with a diameter of 200 nm on a SiN membrane that is sandwiched in a microfluidic cell (see Supplementary Note 1 for details). Figure 1a-c show a schematic description of the key elements and working principle of the system. After the analyte molecules, such as the polypeptides in Figure 1d,e, are adsorbed on the AuNSs (Figure 1c), an electrical potential from 0 to -3 V is applied across the membrane to drive the negatively-charged AuNSs in electrolyte from the lower reservoir to the upper one through one of the gold nanoholes. When the nanohole is resonantly illuminated by a 785 nm laser, it generates an intense plasmonic force that can drag a passing AuNS to its sidewall as near as around 5 nm and trap it for tens of seconds[7]. Before being trapped in the nanohole, the AuNS in solution can have many tips (Figure 1c and Supplementary Figure S1) that act as hot spots with similar field intensities[17]. Once the AuNS is trapped, the plasmonic coupling between the nanohole wall and the AuNS tip near the wall induces a plasmonic gap mode on the AuNS tip with a field intensity at least 2 orders of magnitude higher than those of AuNS tips far from the nanohole wall[18], and ultrahigh localization with a size similar to that of a nucleobase[7]. The SERS signals of the molecules enhanced by the gap mode are so strong that they actually screen SERS signals from other tips, leading to a definite strong, single hot spot (Figure 1b). When analyte molecules with a size similar to the hot spot form submonolayers on the AuNS, the single hot spot covers only 1 such molecule in most cases and thus produce single-molecule SERS spectra (Supplementary Figure S2), as already demonstrated in our previous work[7].

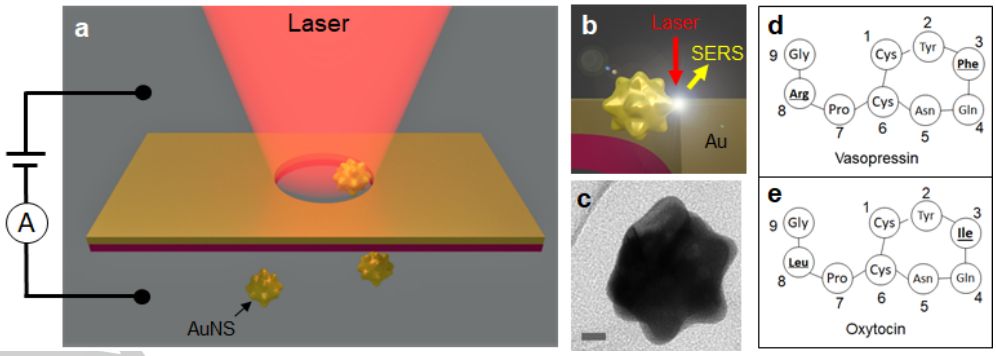


Figure 1. Electro-plasmonic system with single SERS hot spot. (a) Schematic of the microfluidic setup that allows single AuNSs to be trapped in a gold nanohole with plasmonic resonance upon laser excitation at 785 nm. (b) The trapping leads to a single strong and extremely localized SERS hot spot between the AuNS tip and the nanohole sidewall that allows single-molecule SERS. (c) TEM image of the AuNS. The scalebar is 10 nm. The sequences of the 2 polypeptides measured by the electro-plasmonic system: (d) Vasopressin and (e) Oxytocin, in which the different amino acid residues are underlined.

10.1002/anie.202000489

Accordingly, our system produces single-molecule SERS spectra by adsorbing submonolayers of amino acids on the AuNSs (details in Supplementary Notes 1). The single-molecule spectra from a trapped AuNS are strikingly different from multi-molecule spectra produced by the AuNSs diffusing in solution. For example, when sub-monolayers of Arginine (Arg) are adsorbed on the AuNSs (Arg-AuNS), an Arg-AuNS trapped in the nanohole produces stable single-molecule spectra with narrow peaks (Figure 2a and colored curves in Figure 2b). On the contrary, multi-molecule SERS spectra from AuNSs diffusing in solution exhibit irreproducible broad peaks (black curve in Figure 2b and Supplementary Figure S3). The broadening can be due to different conformations of the molecules that are adsorbed on multiple tip hot spots of the diffusing AuNSs. In contrast to an average width of around 15 cm^{-1} for the CH_2 peak at 899 cm^{-1} in the multi-molecule spectra, the single-molecule peaks had an average peak width as small as 7 cm^{-1} (Figure 2c,d). In fact, the broad multi-molecule peak, such as the CH_2 deformation mode at 899 cm^{-1} in Figure 2c, can be regarded as a sum of the shifting single-molecule peaks[19], each of which represented specific conformations and orientations of the adsorbed Arg molecules.[20]

The Raman cross-sections of the non-aromatic amino acids were up to 4 orders of magnitude smaller than those of the aromatic amino acids under resonant excitation at 206.5 nm[21], whereas they remain unknown under non-resonant excitation such as the 785 nm excitation used in our case. To test the sensitivity of the platform, we collect single-molecule spectra of 9 additional different amino acids, including aromatic Phenylalanine (Phe), Tyrosine (Tyr), as well as non-aromatic Glycine (Gly), Leucine (Leu), Isoleucine (Ile), Cysteine (Cys), Proline (Pro), Glutamine (Gln) and Asparagine (Asn) respectively, as reported in Supplementary Note 4. Together with the Arg, they are the building blocks of the OXT and VAS. By considering the surface selection rules for SERS and molecule sizes, the intensities of the 2 strongest single-molecule SERS modes of the 10 amino acids differ within 1 order in magnitude (Table 1). On the other hand, their reproducible single-molecule spectra exhibit similar narrow peaks to those of the Arg (Supplementary Note 4). Successful identification of the amino acids is performed by assigning at least 2 peaks (see Supplementary Table S3). Significantly, all the 8 non-aromatic amino acids are identified at single-molecule level for the first time. Figure 3 lists representative single-molecule spectra of the 10 individual amino acids that show distinct and narrow Raman peaks. They demonstrate an unprecedented spectral discriminative capability of our electro-plasmonic system because of the sensitive single hot spot. They also lay the foundation for using the SERS spectra to discern further single amino acid residues in the OXT and VAS by our system.

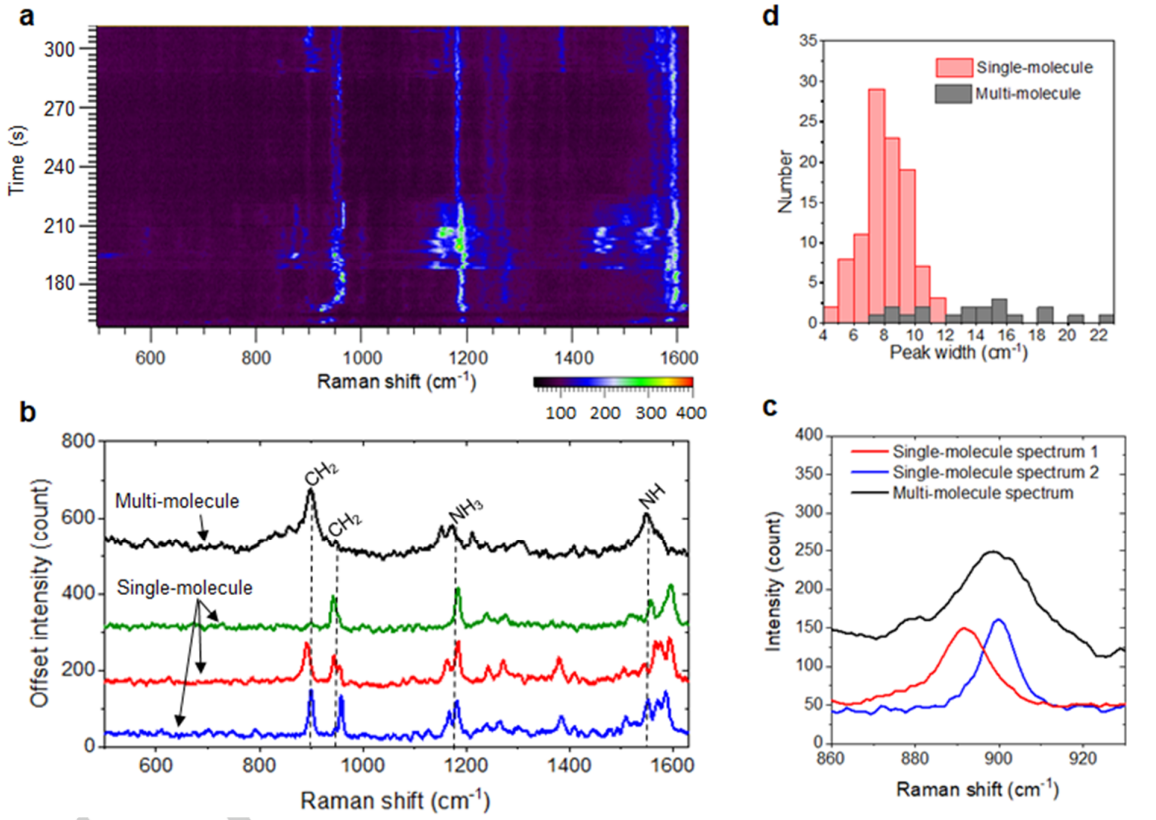


Figure 2. Single-molecule SERS from trapped Arg-AuNSs. (a) SERS time series of trapped submonolayer Arg-AuNS. The color bar represents signal-to-base peak intensity. (b, c) Comparisons between single-molecule spectra (colored lines) of trapped submonolayer Arg-AuNS extracted from (a) and multi-molecule spectrum (black line) from multilayer Arg-AuNS diffusion in solution extracted from Supplementary Figure S3. The vertical black dashed lines indicate assignment of Arg SERS modes according to Ref.[22] (d) Histograms of the 899 cm^{-1} peak widths of the single-molecule spectra (a) and multi-molecule spectra (Supplementary Figure S3), respectively.

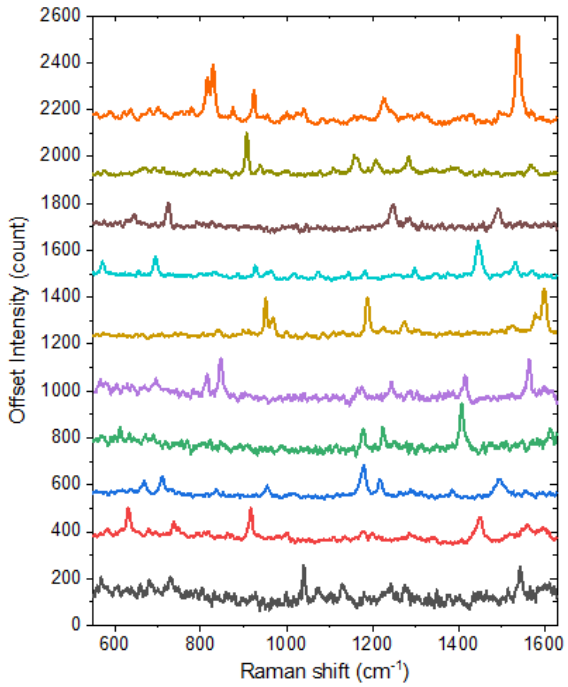


Figure 3. Single-molecule SERS spectra of 10 types of individual amino acids extracted from Supplementary Note 4. Spectra from top to bottom are: Leu, Gln, Asn (intensity × 2), Ile, Arg, Pro (intensity × 3), Phe (intensity × 3), Tyr, Cys, Gly (intensity × 5).

Multiplex discrimination of single amino acid residues in single polypeptides.

Both VAS and OXT are polypeptides of nine amino acids (a nonapeptide). Their sequences are:[16]

VAS: C terminus-Cys1-Tyr2-**Phe3**-Gln4-Asn5-Cys6-Pro7-**Arg8**-Gly9-N terminus

OXT: C terminus-Cys1-Tyr2-**Ile3**-Gln4-Asn5-Cys6-Pro7-**Leu8**-Gly9-N terminus

As underlined, the sequential differences of VAS and OXT are characterized by 2 amino acid residues at position 3 (Phe → Ile) and position 8 (Arg → Leu). Both polypeptides have cycled topology due to S-S bonds formed between the two Cys residues at position 1 and 6, as shown in Figure 1d,e. Since the VAS has the Tyr and Phe residues and the OXT has the Tyr residue, multi-molecule SERS spectra of the polypeptides in literature were usually dominated by the spectra of the corresponding aromatic amino acids.[12, 13]

By adsorbing polypeptide submonolayers on the AuNSs and trapping them in our electro-plasmonic system, SERS spectra of single polypeptides with narrow peaks are produced. In the single-polypeptide spectra, SERS peaks of all the amino acid residues are distinguishable with peak assignments of at least 2 Raman modes. Here we focus on the SERS spectra of the sequence-characterizing residues at position 3 (Phe3 verse Ile3) and position 8 (Arg8 verse Leu8) and the neighboring ones, the spectra of other residues can be found in Supplementary Note 5. For example, the single-VAS SERS spectra show stable Tyr2-only peaks (Figure 4a,b) and Phe3-only peaks (Figure 4c,d) from different measurements, respectively, which last stably for tens of seconds. Similarly, multi-residue peaks are also observed, such as 2-residue peaks of Pro7-Arg8 (Figure 4e,f). The high probability of the multi-residue peaks (Supplementary Note 5) remains even when we lower the number of VAS in the submonolayer adsorbed on the AuNSs (Supplementary Table S1). By identifying the non-aromatic amino acid residues, the residues Phe3 and Arg8, unique to VAS, have all been characterized.

Table 1. Topological polar surface areas and intensities of 2 strongest SERS modes from single-molecule spectra of individual amino acids in Supplementary Note 4.

Amino Acid	Topological polar surface area [23] (Å^2)	SERS modes (cm^{-1})	Intensity (count)
Gly	63.3	1043	55
		1406	43
Cys	64.3	670	62
		1132	28
Leu	63.3	1171	102
		1527	129
Ile	63.3	930	135
		1448	115
Asn	106	724	113
		1249	93
Gln	106	916	252
		1280	102
Arg	128	1180	292
		1595	155
Pro	49.3	844	76
		1565	51
Phe	63.3	812	435
		1199	344
Tyr	83.6	1167	306
		1569	286

Notably, no signals from aromatic amino acid residues appear together with the signals from non-aromatic ones, such as Pro7, in the single-VAS spectra in Figure 4e, even though SERS modes of individual aromatic Phe or Tyr are around 8 times in intensity stronger than those of Pro from Table 1. It means that the single-residue peaks in Figure 4a-d are due to excitation of one aromatic amino acid residue in the VAS chain. Furthermore, the peaks of different residues of single VAS in Figure 4 suggest that the hot spot covers and excites different segment of the VAS chain rather than the whole VAS molecule.

The AuNSs in this work have an average diameter of 50 nm and their tips have an average apex diameter of around 10 nm (Figure 1c). The extremely localized hot spot here should be due to random atomic protrusions on the AuNS tip, which were reported in recent SERS studies of gold nanoparticles on gold films [20, 24] as well as in literatures of tip-enhanced Raman spectroscopy[25] by other groups. Such a small hot spot was also observed in our previous work that detected single nucleobases in single oligonucleotide molecules.[7] Here the size of the hot spots can be estimated by the size of single Tyr molecule between

10.1002/anie.202000489

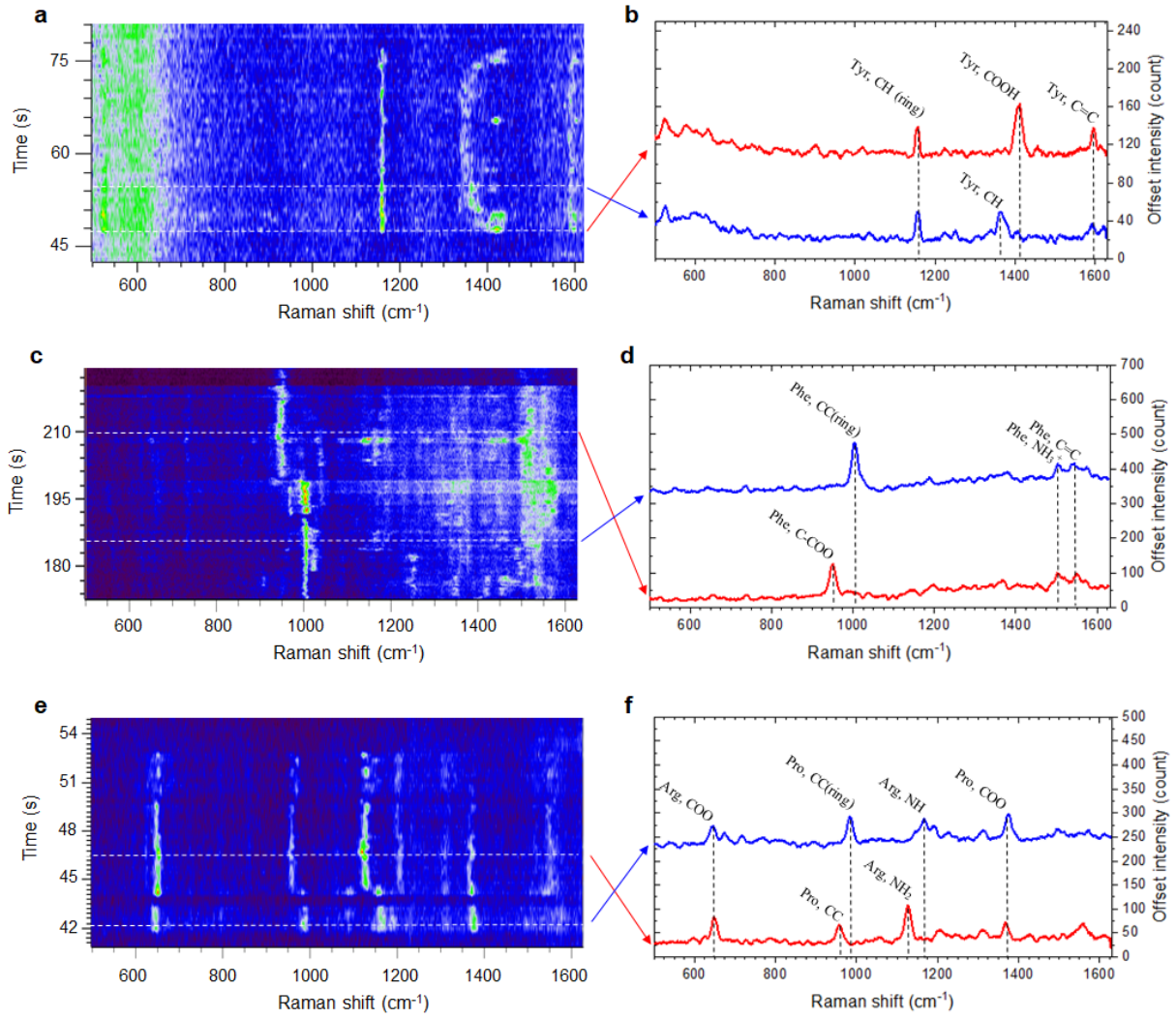


Figure 4. Multiplex discrimination of amino acid residue in single Vassopresin adsorbed on the AuNS (VAS-AuNS). SERS time series of (a,b) single Tyr2, (c,d) single Phe3 and (e,f) Pro7-Arg8 of different single VAS-AuNSs, respectively. Colored arrows indicate the spectra in (b,d,f) corresponding to times in (a,c,e) marked by the white dotted lines. Black dotted lines in (b,d,f) indicate frequency positions of the SERS modes of corresponding amino acids according to Refs.[22, 26-28]

the topological polar surface area of Tyr (83.6 Å^2 from Table 1) and a maximum accessible surface area of Tyr (255 Å^2 from Supplementary Table S1). The topological polar surface area and the intensities of the SERS modes of individual Arg are comparable to those of Tyr. Yet single-Arg8 peaks are not observed which might be a result of a folded conformation of the sidechain.

A common feature shared by the multiplex spectra in Figure 4 is that some SERS modes blink while some remain constant, suggesting changes of molecular conformations[29]. Such features are also found in some spectra of individual amino acids (Figure 2a and Supplementary Note 4). In particular, the absence of ring breath modes of Phe, Tyr and Pro can be used to characterize geometry of the molecules on the surface due to the surface selection rules for SERS[30]. Briefly, the plasmonic hot spot on the AuNS tip has electric polarization parallel to the surface normal of the tip surface[7]. If an aromatic ring is flat on the tip surface, its ring breathing vibration is vertical to the hot spot polarization and thus can not be enhanced by the hot spot, leading to suppressed intensity of the ring breath mode[28]. When the aromatic ring is tilted with respect to the tip surface, its ring breathing vibration has a component that is parallel to the hot spot polarization and thus can be enhanced by the hot spot [30]. Accordingly, the absence of ring breath doublets at 860 and 828 cm^{-1} of Tyr characterize a flat adsorption geometry on surface[28], which is the case of Figure 4a,b. Also, the absence of the Phe ring breath mode, CC(ring), at 1006 cm^{-1} characterize a flat adsorption geometry of Phe3[28]. Therefore, Figure 4c,d presents a rotation process of the Phe3 ring: the Phe3 ring is tilted at 174 s due to strong ring breath mode at 1006 cm^{-1} and then turns flat when the ring breath mode vanishes at around 201 s. Moreover, the absence of the Pro ring

10.1002/anie.202000489

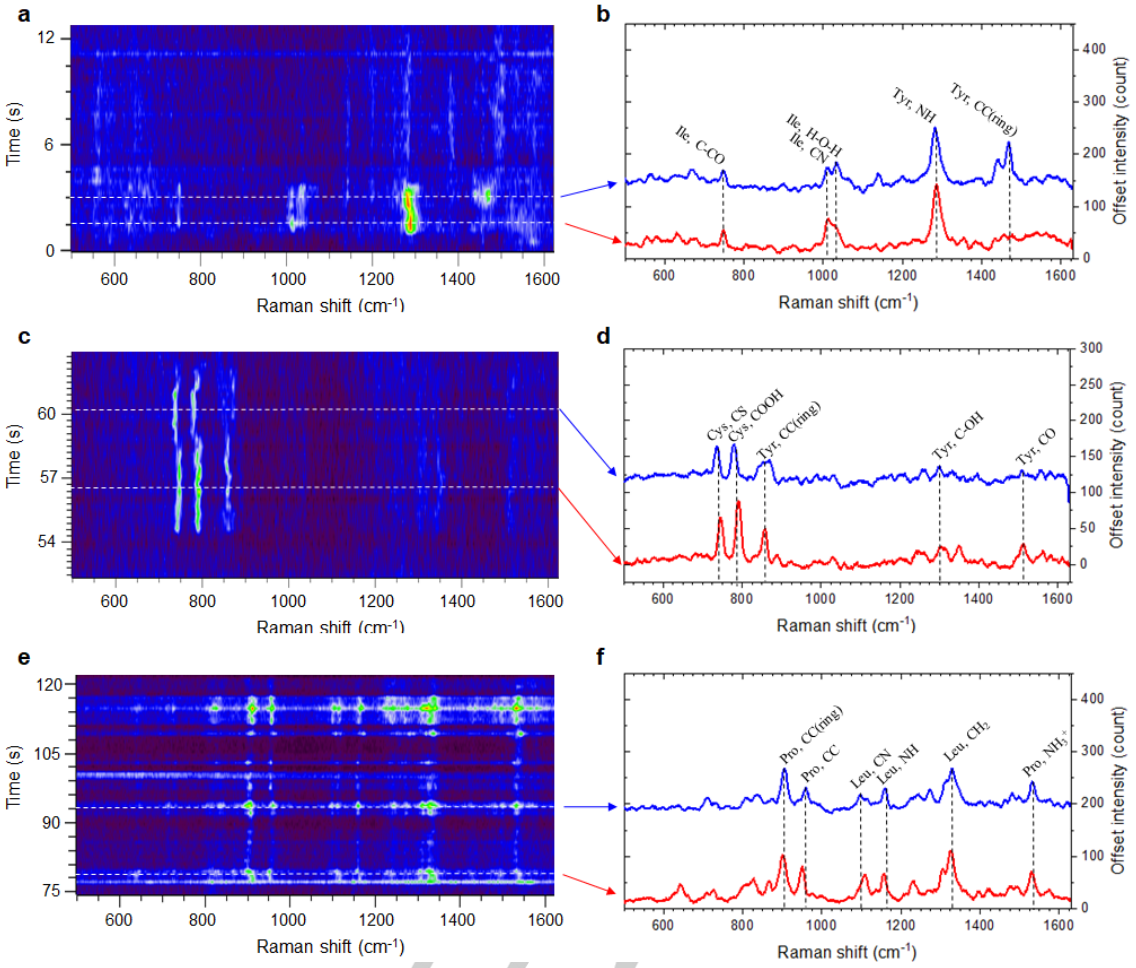


Figure 5. Multiplex discrimination of amino acid residue in single Oxytocin adsorbed on the AuNS (OXT-AuNS). Multiplex SERS time series of (a,b) Tyr2-Ile3, (c,d) Cys1-Tyr2, and (e,f) Pro7-Leu8 of different single OXT-AuNSs, respectively. Colored arrows indicate the correspondence of the spectra in (b,d,f) corresponding to times in (a,c,e) marked by the white dotted lines. Black dotted lines in (b,d,f) indicate frequency positions of the SERS modes of corresponding amino acids according to Refs.[22, 26-28]

modes, CC(ring), at 990 and 910 cm^{-1} suggests flat pyrrolidine ring of Pro on the surface[26], which is the case at 46.5 s in Figure 4e,f. A rotation of the Pro7 pyrrolidine ring from tilted to flat geometry due to the disappearance of the ring mode of 990 cm^{-1} occurs at 43.5 s.

SERS spectra of single OXT (Figure 5) are also measured by adsorbing OXT submonolayers on the AuNSs and trapping them in the electro-plasmonic platform. Different from the single-VAS case, no single-residue peaks are observed in the single-OXT spectra. Significantly, Tyr2-Ile3 peaks are observed for a few seconds in Figure 5a,b as the only case that signals of both aromatic and non-aromatic amino acid residues appear together. The Tyr2 modes have intensities around twice those of Ile3 modes. Tyr2 is in flat geometry due to absence of the ring doublets at 860 and 828 cm^{-1}.[28] Conformational change of Tyr2 is more obvious in the Cys1-Tyr2 peaks in Figure 5c,d where the Tyr2 ring turns to be more tilted on the gold surface during 56 - 58.5 s due to strengthening of the ring mode at 860 cm^{-1}. Pro7-Leu8 peaks in Figure 5e,f exhibit moderate intensity of the Pro ring mode at 910 cm^{-1} to suggest a tilted Pro7 ring [26]. As expected, the amino acid residues (Ile3 and Leu8) unique to OXT have all been characterized by our system. Direct comparison between Figure 4 and Figure 5 (summarized in Supplementary Figure S12) identifies clearly the residue differences between VAS and OXT at position 3 (Phe3 verse Ile3) and position 8 (Arg8 verse Leu8). It is the narrow peak width that allows the multiplex discrimination.

Influence of the hot spot to the sidechain conformations.
Although the ring breath modes can characterize the ring geometry of the aromatic amino acid residues, the conformations of the non-aromatic ones remain difficult to explain by the SERS spectra. To have a better understanding of the sidechain conformations, we use Molecule Dynamics (MD) to simulate the VAS and OXT on gold surfaces in an aqueous environment, respectively (details in Supplementary Note 6). Figure 6a, b show 2 most probable conformations of the VAS on gold surfaces,

10.1002/anie.202000489

some residues of which are in agreement with the spectra in Figure 4. Briefly, the rings of Tyr2, Phe3 and Pro7 residues are flat on the surface in Figure 6a, whereas Figure 6b shows a tilted Phe3 ring and vertical Pro7 ring. Similarly, MD models of single OXT on gold surface (Figure 6c, d) are in agreement with the aromatic sidechain conformations derived from the SERS spectra in Figure 5. In the model in Figure 6c, Tyr2 ring remains flat on the surface and interacts with the Ile3, which may explain the Tyr2-Ile3 modes in Figure 5a,b. In addition, the Cys1-Cys6 bond is attached on the gold surface, in agreement with the Cys1-Tyr2 modes (Figure 5c,d). In both Figure 6c and 6d, the Pro7 ring is tilted, in agreement with Pro7-Leu8 modes in Figure 5e,f. Therefore, the correlations between the simulated aromatic ring geometries and those derived from the SERS spectra justify our MD models. The single hot spots due to the random atomic protrusions on the AuNS tip could be localized at the positions of these residues to produce the corresponding spectra.

The MD-based models provide insight to the SERS features of other non-aromatic amino acid residues. For instance, very few Cys peaks are observed in the VAS spectra (Supplementary Figure S10), probably because the Cys1-Cys6 are distant from the gold surface according to Figure 6a. Furthermore, the cycled structure due to the Cys1-Cys6 bond limits the conformation patterns but leaves more freedom to the Pro7-Arg8-Gly9 tail, which are affected by the pyrrolidine ring of Pro7. The Arg8 is close to Gly9 (Figure 6a) when the pyrrolidine ring of Pro7 is flat on the surface. The Arg8 is far from Gly9 (Figure 6b) when the pyrrolidine ring of Pro7 is vertical on the surface. Here, we can see from the models that the rings of the aromatic amino acid residues tend to touch the gold surface as low-energy conformations (Supplementary Note 6) and they are thus enhanced the most by the hot spot. They can prevent other non-aromatic amino acid residues from approaching the enhancing surface due to their rigid structures and large spatial occupancy[31]. This may explain the dominance of the SERS peaks of the aromatic amino acid residues in multi-molecule protein SERS spectra in literature. Therefore, it is the small hot spot of our system that leads to the localized excitation and discrimination of the non-aromatic amino acid residues. Otherwise, if a hot spot is large enough to cover the whole VAS molecule, the aromatic amino acid residues will be enhanced more than the non-aromatic ones, leading to the dominance of those signals.

From the MD simulations, the polypeptides actually change their conformations on the gold surface from time to time (Supplementary Note 6). The dynamics of the sidechain peaks in the SERS time series of single polypeptides in Figure 4, 5 and Supplementary Note 5 suggest that the polypeptide chains move across the single hot spots. The polypeptide conformation can be influenced by the strong hot spot that can induce two possible effects: heating[20] and optical trapping[32]. The former increases the thermal diffusion of the molecule, whereas the latter traps the molecule in the hot spot. For individual aromatic amino acids, such as Phe and Tyr, their SERS spectra suggest that their aromatic rings are kept flat in the hot spot for ≥ 60 s (Supplementary Figure S8). In comparison, another single-molecule study at low temperature indicated that a dye molecule with aromatic rings took <10 s to rotate on gold surface at 90 K[29]. The room temperature in our measurement of individual Phe and Tyr should speed up the rotations. Therefore, the difference of rotational times suggests a dominant role of the optical trapping effect in the case of individual amino acids. However, when it comes to polypeptides, the rotational times of the aromatic rings of their aromatic sidechains become shorter, ranging from 1 to 30 s, as suggested by the spectra in Figure 4. It is obvious that the molecule size also plays a role, and more work is needed to investigate it. These observations demonstrate that our electro-plasmonic system has the potential to study conformation changes that large biomolecules, such as polypeptides, can undergo on gold surfaces in aqueous environment.

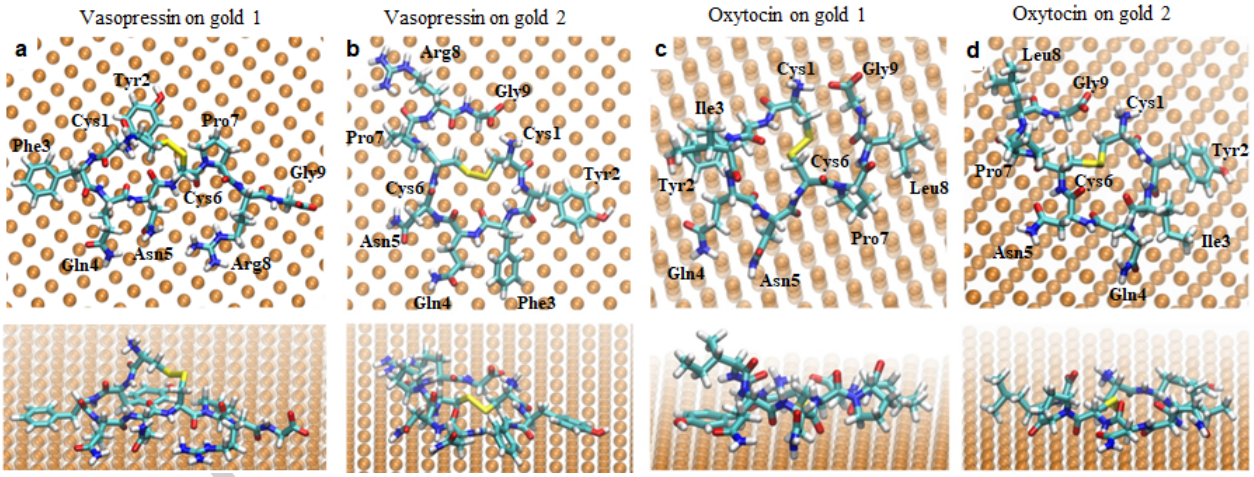


Figure 6. Representative structures of the most probable absorbed conformations of the polypeptides on a gold surface obtained via molecular dynamics simulations. (a,b) 2 most probable conformations for Vasopressin on a gold surface. (c,d) 2 most probable conformations for Oxytocin on a gold surface. Upper panels are the top views of the molecule while lower panels are the side views.

Conclusion

By trapping gold nanoparticles with adsorbed submonolayers of molecules to generate single strong hot spots, our electro-plasmonic system demonstrates multiplex SERS spectroscopy with narrow peaks and single-molecule sensitivity. 10 types of amino acids with and without aromatic rings are firstly measured at the single-molecule level to demonstrate the sensitivity and spectral discriminative capability. Then, by measuring single VAS and OXT individually, the produced multiplex spectra not only discriminate the residue differences between them, but also characterize conformation dynamics of the polypeptides on gold surfaces with corresponding simulations by Molecule Dynamics. The Molecule Dynamics models in our work provide insight into the problem of how the spatial occupancy of aromatic rings on surfaces leads to the signal dominance of aromatic amino acids residues in protein SERS.

By constructing a single hot spot with single-molecule sensitivity and high localization, our work solves the problem that signals of non-aromatic amino acid residues are invisible in protein SERS spectra, laying a foundation for the single-molecule protein sequencing. In fact, the 3-residue peaks from the single-vasopressin spectra can discriminate 3 different amino acid residues in a single polypeptide chain consisting of 9 residues, which suggests a sequencing capacity of the protein sequences up to 3^9 = 19,683.[33] Recently, an ultrafast SERS technique allowed detection time of single-molecule events to reach microsecond scale[34]. A combination of our platform and the ultrafast detection technique can pay the way to single-protein detection and analysis. Therefore, we believe that our work represents a significant improvement towards single-molecule protein sequencing.

Supporting Information

Details of materials, fabrications, Raman measurement, characterization, data processing, Molecular Dynamics simulations and additional SERS spectra can be found in Supporting Information.

Acknowledgements

The authors thank Dr. Michele Dipalo and Dr. Daniel Darvill for drawing of schematic figures and English editing respectively. Walter Rocchia thanks Dr. Stefano Corni for providing the topology of the gold substrate suitable for the use of the GOLP MD force field. The research leading to these results has received funding from the Horizon 2020 Program, FET-Open: PROSEQO, Grant Agreement no. [687089].

Author contributions

F.D.A., D.G., J.A.H. and M.Z.M. conceived the project. J.A.H carried out Raman experiments and wrote the manuscript. M.Z.M., G.G. and Y.Z. handled the molecules' attachment on the AuNSs and the AuNS stability. G.G. and Y.Z. measured the Zeta potentials and absorbance of the functioned AuNSs. Y.Z. fabricated the nanohole device. M.A.S and W.R. did the Molecular Dynamics simulation. J.A.H. and A.H. analyzed the data. F.D.A. and D.G. supervised the work. All authors discussed the results and contributed to the final manuscript preparation.

Keywords: Single molecule• Raman spectroscopy • Amino acid • Protein • Molecular Dynamics

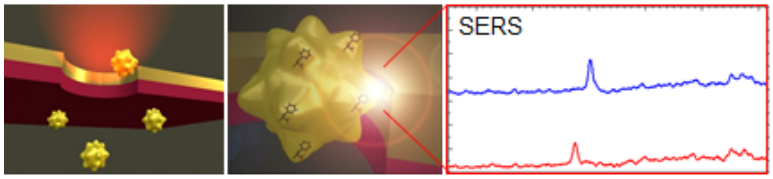
[1] J. Shendure, S. Balasubramanian, G. M. Church, W. Gilbert, J. Rogers, J. A. Schloss, R. H. Waterston, *Nature* **2017**, *550*, 345-353.
[2] D. Deamer, M. Akeson, D. Branton, *Nature Biotechnology* **2016**, *34*, 518-524.
[3] L. Restrepo-Perez, C. Joo, C. Dekker, *Nature Nanotechnology* **2018**, *13*, 786-796.
[4] J. P. da Costa, P. S. M. Santos, R. Vitorino, T. Rocha-Santos, A. C. Duarte, *Trac-Trends in Analytical Chemistry* **2017**, *93*, 171-182.
[5] J. van Ginkel, M. Filius, M. Szczepaniak, P. Tulinski, A. S. Meyer, C. Joo, *Proceedings of the National Academy of Sciences of the United States of America* **2018**, *115*, 3338-3343; O. N. Assad, T. Gilboa, J. Spitzberg, M. Juhasz, E. Weinhold, A. Meller, *Advanced Materials* **2017**, *29*.
[6] C. Chen, Y. Li, S. Kerman, P. Neutens, K. Willems, S. Cornelissen, L. Lagae, T. Stakenborg, P. Van Dorpe, *Nature Communications* **2018**, *9*.
[7] J.-A. Huang, M. Z. Mousavi, Y. Zhao, A. Hubarevich, F. Omeis, G. Giovannini, M. Schütte, D. Garoli, F. De Angelis, *Nature Communications* **2019**, *10*, 5321.
[8] J. D. Spitzberg, A. Zrehen, X. F. van Kooten, A. Meller, *Advanced Materials* **2019**, *31*; D. Garoli, H. Yamazaki, N. Maccaferri, M. Wanunu, *Nano Letters* **2019**; D. Verschueren, X. Shi, C. Dekker, *Small Methods* **2019**, *3*.
[9] S. E. J. Bell, N. M. S. Sirimuthu, *Chemical Society Reviews* **2008**, *37*, 1012-1024.
[10] Y. Zhao, J.-A. Huang, Z. Zhang, X. Chen, W. Zhang, *Journal of Materials Chemistry A* **2014**, *2*, 10218-10224; S. Tian, O. Neumann, M. J. McClain, X. Yang, L. Zhou, C. Zhang, P. Nordlander, N. J. Halas, *Nano Letters* **2017**, *17*, 5071-5077.
[11] P. Matteini, M. Cottat, F. Tavanti, E. Panfilova, M. Scuderi, G. Nicotra, M. C. Menziani, N. Khlebtsov, M. de Angelis, R. Pini, *Acs Nano* **2017**, *11*, 918-926.
[12] E. Podstawka, Y. Ozaki, L. M. Proniewicz, *Applied Spectroscopy* **2004**, *58*, 1147-1156.
[13] E. Podstawka, E. Sikorska, L. M. Proniewicz, B. Lammek, *Biopolymers* **2006**, *83*, 193-203.
[14] F. Wei, D. Zhang, N. J. Halas, J. D. Hartgerink, *Journal of Physical Chemistry B* **2008**, *112*, 9158-9164.
[15] L. J. Xu, C. Zong, X. S. Zheng, P. Hu, J. M. Feng, B. Ren, *Analytical Chemistry* **2014**, *86*, 2238-2245.
[16] N. Japundzic-Zigon, *Current Neuropharmacology* **2013**, *11*, 218-230.
[17] D. Issaad, H. Moustaoui, A. Medjahed, L. Lalaoui, J. Spadavecchia, M. Bouafia, M. L. de la Chapelle, N. Djaker, *Journal of Physical Chemistry C* **2017**, *121*, 18254-18262; W. X. Niu, Y. A. A. Chua, W. Q. Zhang, H. J. Huang, X. M. Lu, *Journal of the American Chemical Society* **2015**, *137*, 10460-10463.
[18] E. Petryayeva, U. J. Krull, *Analytica Chimica Acta* **2011**, *706*, 8-24; J. J. Baumberg, J. Aizpurua, M. H. Mikkelsen, D. R. Smith, *Nature Materials* **2019**, *18*, 668-678.
[19] P. G. Etchegoin, E. C. Le Ru, *Analytical Chemistry* **2010**, *82*, 2888-2892.
[20] F. Benz, M. K. Schmidt, A. Dreismann, R. Chikkaraddy, Y. Zhang, A. Demetriadou, C. Carnegie, H. Ohadi, B. de Nijs, R. Esteban, J. Aizpurua, J. J. Baumberg, *Science* **2016**, *354*, 726-729.
[21] Z. H. Chi, X. G. Chen, J. S. W. Holtz, S. A. Asher, *Biochemistry* **1998**, *37*, 2854-2864.
[22] A. E. Aliaga, C. Garrido, P. Leyton, G. Diaz, J. S. Gomez-Jeria, T. Aguayo, E. Clavijo, M. M. Campos-Vallette, S. Sanchez-Cortes, *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy* **2010**, *76*, 458-463.
[23] S. Kim, J. Chen, T. J. Cheng, A. Gindulyte, J. He, S. Q. He, Q. L. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, E. E. Bolton, *Nucleic Acids Research* **2019**, *47*, D1102-D1109.
[24] H.-H. Shin, G. J. Yeon, H.-K. Choi, S. M. Park, K. S. Lee, Z. H. Kim, *Nano Letters* **2018**, *18*, 262-271.
[25] R. Zhang, Y. Zhang, Z. C. Dong, S. Jiang, C. Zhang, L. G. Chen, L. Zhang, Y. Liao, J. Aizpurua, Y. Luo, J. L. Yang, J. G. Hou, *Nature* **2013**, *498*, 82-86; Z. He, Z. Han, M. Kizer, R. J. Linhardt, X. Wang, A. M. Sinyukov, J. Wang, V. Deckert, A. V. Sokolov, J. Hu, M. O. Scully, *Journal of the American Chemical Society* **2019**, *141*, 753-757; P. Liu, D. V. Chulhai, L. Jensen, *Acs Nano* **2017**, *11*, 5094-5102.
[26] J. J. Carcamo, A. E. Aliaga, E. Clavijo, C. Garrido, J. S. Gomez-Jeria, M. M. Campos-Vallette, *Journal of Raman Spectroscopy* **2012**, *43*, 750-755; Y. S. Mary, L. Ushakumari, B. Harikumar, H. T. Varghese, C. Y. Panicker, *Journal of the Iranian Chemical Society* **2009**, *6*, 138-144.
[27] E. Podstawka, Y. Ozaki, L. M. Proniewicz, *Applied Spectroscopy* **2005**, *59*, 1516-1526; T. Deckert-Gaudig, V. Deckert, *Small* **2009**, *5*, 432-436.
[28] T. Deckert-Gaudig, E. Rauls, V. Deckert, *Journal of Physical Chemistry C* **2010**, *114*, 7412-7420.
[29] K.-D. Park, E. A. Muller, V. Kravtsov, P. M. Sass, J. Dreyer, J. M. Atkin, M. B. Raschke, *Nano Letters* **2016**, *16*, 479-487.
[30] M. Moskovits, J. S. Suh, *Journal of Physical Chemistry* **1984**, *88*, 5526-5530; M. Moskovits, J. S. Suh, *Journal of Physical Chemistry* **1988**, *92*, 6327-6329.
[31] D. Kurouski, T. Postiglione, T. Deckert-Gaudig, V. Deckert, I. K. Lednev, *Analyst* **2013**, *138*, 1665-1673.
[32] M. Belkin, S.-H. Chao, M. P. Jonsson, C. Dekker, A. Aksimentiev, *Acs Nano* **2015**, *9*, 10598-10611.
[33] Y. Yao, M. Docter, J. van Ginkel, D. de Ridder, C. Joo, *Physical Biology* **2015**, *12*.
[34] N. C. Lindquist, C. D. L. de Albuquerque, R. G. Sobral-Filho, I. Paci, A. G. Brolo, *Nature Nanotechnology* **2019**, *14*, 981-987.

10.1002/anie.202000489

WILEY-VCH

RESEARCH ARTICLE

Entry for the Table of Contents



Sensing of protein sequences by surface-enhanced Raman spectroscopy (SERS) suffers from the signal dominance of aromatic amino acid residues and backbones, which hampers the ability to detect nonaromatic amino acid residues. Here, we use a single, localized SERS hot spot in an electroplasmonic trap that allows for detection of aromatic and nonaromatic residues in single polypeptides due to the hot spot close in size to that of an amino acid.

Institute and/or researcher Twitter usernames: @IITalk and @OProseq