Received: 13 April 2020 | Revised: 26 May 2020 | Accepted: 5 June 2020

DOI: 10.1002/qua.26369

FULL PAPER

International Journal of Quantum Chemistry WILEY

The conformational space of a flexible amino acid at metallic surfaces

Dmitrii Maksimov[1,2] | Carsten Baldauf[1] | Mariana Rossi[1,2]

[1]Fritz Haber Institute of the Max Planck Society, Berlin, Germany

[2]Max Planck Institute for Structure and Dynamics of Matter, Hamburg, Germany

Correspondence
Mariana Rossi, Max Planck Institute for Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany.
Email: mariana.rossi@mpsd.mpg.de

Funding information
BigMAX; Max Planck - EPFL Center for Molecular Nanoscience and Nanotechnology

Abstract

In interfaces between inorganic and biological materials relevant for technological applications, the general challenge of structure determination is exacerbated by the high flexibility of bioorganic components, chemical bonding, and charge rearrangement at the interface. In this paper, we investigate a chemically complex building block, namely, the arginine (Arg) amino acid interfaced with Cu, Ag, and Au (111) surfaces. We investigate how the environment changes the accessible conformational space of this amino acid by building and analyzing a database of thousands of structures optimized with the Perdew-Burke-Ernzerhof (PBE) functional, including screened pairwise van der Waals interactions. When in contact with metallic surfaces, the accessible space for Arg is dramatically reduced, while the one for Arg-H^+ is instead increased if compared to the gas phase. This is explained by the formation of strong bonds between Arg and the surfaces and by their absence and charge screening on Arg-H^+ upon adsorption. We also observe protonation-dependent stereoselective binding of the amino acid to the metal surfaces: Arg adsorbs with its chiral $C_\alpha H$ center pointing H away from the surfaces, while Arg-H^+ adsorbs with H pointing toward the surface.

KEYWORDS

adsorption, arginine, conformational space, electronic structure, metal, DFT

1 | INTRODUCTION

The organic-inorganic interfaces formed between peptides and surfaces are of interest to diverse fields such as medicine, optoelectronics, and energy storage.[1–10] Amino acids and their oligomers—that is, peptides—are particularly interesting materials because they are naturally biocompatible and offer a rich functional space already at the amino acid level, which can be extended by the combinatorial increase of molecular motifs available through the peptide bond formation. In these setups, the inorganic component offers a platform to immobilize and template the bioorganic counterpart, as well as to record electronic signals from interactions or reactions. However, emergent behavior makes it impossible to base the description of such interfaces solely on the study of its independent components.

On the experimental front, progress on the fundamental understanding of these interfaces has been achieved with recent soft-landing techniques[11] that allow amino acids and peptides of a few tens of building blocks to be deposited on solid surfaces and subsequently imaged in ultra-high vacuum conditions using scanning tunneling microscopy (STM).[12] These experiments have shown that small sequence modifications,

changes in the protonation of a peptide, and charge transfer at the interface can yield drastic changes in the two-dimensional structure and self-assembly.[11,13,14]

Explaining the physical mechanisms that govern adsorption and pattern formation calls for systematic theoretical investigations with an atomistic resolution of such interfaces. However, such studies are challenging as they require (a) accurate energetics for a system build-up of elements across the periodic table and where considerable charge rearrangement and chemical reactions can occur, (b) sampling and representing a large conformational space, and (c) dealing with structure motifs that can only be represented by unit cells containing hundreds of atoms. Pioneering work that used density functional theory (DFT) to study amino acids at inorganic substrates has focused on small or rigid amino acids and a limited portion of their conformational space.[15–19] One of the first studies that were dedicated to larger amino acids was performed by Hong and coworkers[20] and illustrates the challenge of properly sampling the large structure space of flexible biomolecules. These studies have clarified that an accurate energy function is only one of the ingredients needed to properly predict the structure of peptides at surfaces, with the sampling of structure space being just as important.

We propose to analyze how the structure space of a complex building block with multiple functional groups is affected by changes in its (non-biological) environment. To achieve this goal, we analyze a database of the arginine (Arg) amino acid and its protonated counterpart (Arg-H^+) in the gas phase and interfaced with Cu(111), Ag(111), and Au(111). Arginine is a good testbed because it is a relatively small molecule but contains a very flexible side chain and allows for different stable protonation states. We show that it is possible to perform an exhaustive conformational space exploration while still treating the potential energy surface (PES) with DFT, thus capturing the charge rearrangements at the interface. The outcome of such a computational structure search is a large number of stationary states on the respective PES. Not only the global minimum is of interest, but also the relative positioning of other local minima with respect to the structural degrees of freedom are relevant because they can reveal different basins and structures of interest under different conditions.[21–25] In order to handle such high-dimensional data, we make use of recent advances in machine-learning methods that can help to visualize conformational preferences of adsorbed molecules.[26–31]

In the following, we discuss the impact of the protonation state and the presence or absence of the surface on the accessible conformational space of arginine. We start by describing the procedure we followed to build the database of Arg and Arg-H^+ on Cu(111), Ag(111), and Au(111), based on thousands of first-principles structure optimizations that we make available to the community.[32] We then analyze this database and show how different patterns of bonding and charge transfer induce fundamental changes in the accessible conformational space. We also provide an analysis of property trends across the different metallic surfaces, including protonation-dependent stereoselective binding to the surfaces and deprotonation propensities.

2 | COMPUTATIONAL METHODS

Arginine is an amino acid with a flexible side chain of three (aliphatic) CH_2 groups and a guanidino group. A depiction of the Arg molecule, including the labeling of the different chemical groups and specific atoms we will refer to in the manuscript, is shown in Figure 1A. In the context of this publication, we use the term *protonation state* to distinguish between Arg and its single-protonated form Arg-H^+. We use the word *protomers* to distinguish different arrangements of protons within molecules of the same sum formula, for example, the protomers **P1** to **P5** of Arg or the protomers **P6** and **P7** of Arg-H^+, shown in Figure 1B and C. In this section, we describe the computational setup, convergence tests, and sampling strategies we used in order to build the database of Arg and Arg-H^+ adsorbed on different metal surfaces.

All the electronic structure calculations were carried out using the numeric atom-centered basis set all-electron code FHI-aims.[33,34] We used the standard *light* settings of FHI-aims for all species, except when stated otherwise. For modeling the adsorbed molecules, a 5×6 surface unit cell with $4 \times 4 \times 1$ *k*-point sampling was used. The slab contains four layers, and we added a 50 Å vacuum in the *z*-direction in order to separate periodic images of the system. A surface unit cell of this size does not completely isolate neighboring molecules on the surface plane. In order to estimate the magnitude of this spurious interaction, we calculated binding energies for three Arg and three Arg-H^+ structures adsorbed on Cu(111) using different surface unit cell sizes. As shown in the supporting information (SI) (Tables S3 and S4), the relative binding energies change by no more than 50 meV when reaching a 10×12 cell.

An accurate description of the dispersion interactions that play a role in weakly bonded adsorbates on metallic surfaces can be achieved with newly proposed methods that take into account electronic screening and the many-body nature of the dispersion term[35]; however, such methods are not yet computationally feasible for large-scale studies. Considering dispersion interactions is, however, crucial to model such interfaces.[36–43] We thus use the PBE exchange-correlation functional augmented by the TS-vdW^{surf}[36] dispersion correction between pairs of atoms that involve the adsorbed molecule (molecule-molecule and molecule-surface dispersion). Adding pairwise vdW corrections for pairs of metal atoms did not result in a systematic improvement of the lattice constants due to the complex many-body nature of the electronic screening within the metallic bulk. Because the PBE lattice constants for Cu, Ag, and Au are in good agreement with experimental data (Table S1), and the electronic structure itself is not changed by the inclusion of these types of vdW interactions, we chose to use the simplest setup. We optimized all structures until all forces in the system were below 0.01 eV/Å. We also fixed the two bottom layers of the slabs in all optimizations. A dipole correction was applied in the *z*-direction to compensate for the dipole formed by the asymmetric surface configurations.

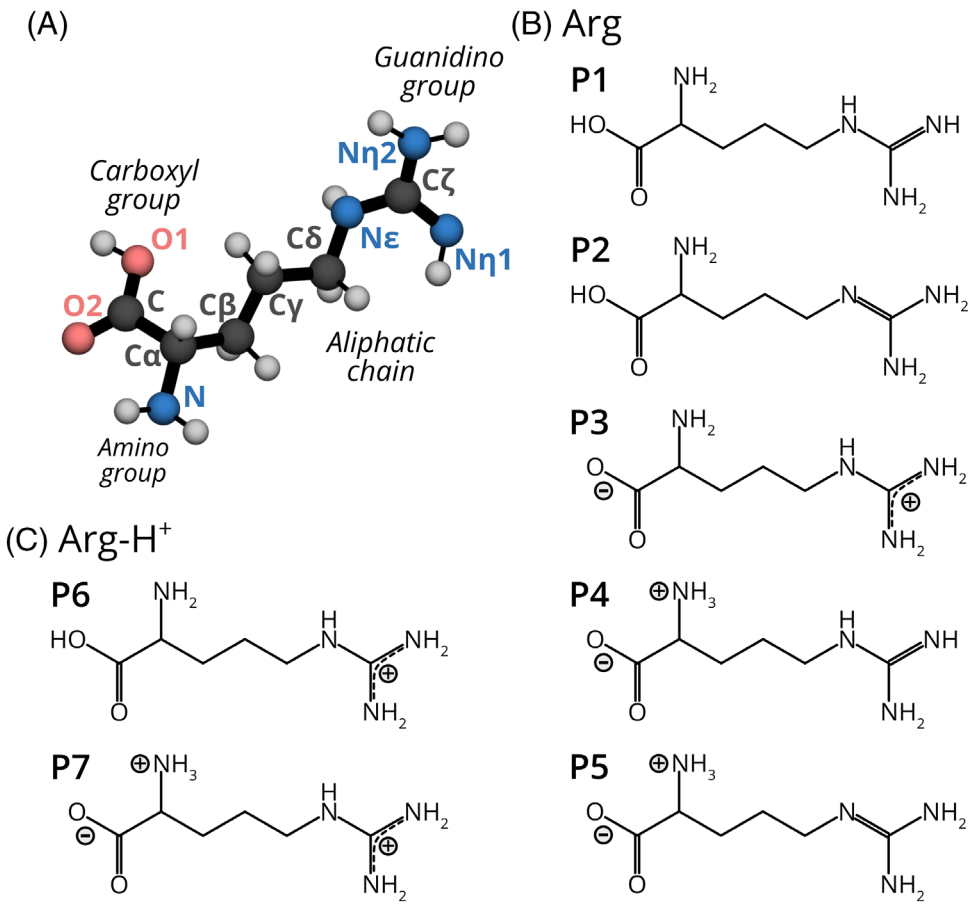


FIGURE 1 A, Pictorial representation of the arginine amino acid, including labels of chemical groups and atoms. B, Protomers of Arg that are addressed in this work. C, Protomers of Arg-H^+ that are addressed in this work

TABLE 1 Number of calculated Arg and Arg-H^+ structures in isolation and adsorbed on Cu(111), Ag(111), and Au(111)

	Gas phase	Cu(111)	Ag(111)	Au(111)
Arg	1206	1156	327	209
Arg-H^+	215	914	718	721

2.1 | Database generation

The sampling of the structure space of the amino acid in two protonation states on metallic surfaces was performed by starting from a previously published dataset comprising stationary points of isolated amino acids and dipeptides.[21,44] For Arg, 1206 structures were present in the database. In order to reduce the number of possibilities, but keeping a representative share of the structures, we considered the 300 lowest energy conformers, the 27 highest energy conformers, and 125 conformers uniformly spanning the energy range in between. For the Arg-H^+ amino acid, all 215 structures present in the gas-phase dataset were used in this study.

We distinguish *upstanding* positions of the molecules, where the largest eigenvector of the rigid-body moment of inertia tensor is approximately perpendicular to the surface plane, from *flat-lying* positions with an arrangement parallel to the surface. For Arg, three lying configurations per structure were generated by randomly placing the molecule flat on the Cu(111) surface and then rotating it by 120° around the principal axis. Two upstanding configurations were generated for the 25 lowest-energy gas-phase structures by first placing the molecule in a random upright orientation and then flipping it. For Arg-H^+, a similar procedure was adopted: flat-lying positions were created by 90° rotations around the principal axis, and upstanding configurations were created for 27 structures that were uniformly selected from the wholeenergy range. In summary, we considered a total of 1156 conformers of Arg@Cu(111) and 914 conformers of Arg-H^+@Cu(111).

All optimized structures that fell within a range of 0.5 eV from the global minimum on Cu(111) were transferred to Ag(111) and Au(111) and further optimized. In addition, we randomly picked 105 Arg-H^+ structures representing the higher-energy range on Cu(111) to be further optimized on Ag(111) and Au(111). Moreover, for Arg, 180 randomly picked structures representing the higher-energy range were considered on Ag(111) and 61 on Au(111). The total amount of calculated structures for each case is summarized in Table 1.

We checked that this strategy ensured a sufficient sampling of the low-energy range of both Arg and Arg-H^+ on Ag(111) and Au(111) by analyzing the alterations in relative energy hierarchies on the different surfaces. In Figure 2A-D, each dot corresponds to a conformer that was optimized first on the Cu(111) surface and then postrelaxed on Ag(111) or Au(111). Within the lowest 0.5 eV range, we do not observe any significant rearrangement of the energy hierarchy with respect to the Cu(111) surface. The energy hierarchies of both Arg and Arg-H^+ on the Ag(111) and Au(111) surfaces are almost identical. The most pronounced outliers in all plots correlate with a higher root mean square displacement (RMSD) of the molecular atoms (that is, disregarding the surface-adsorption site), thus pointing to a structural rearrangement of the molecule.

2.2 | Structure space representation and analysis

We analyze the structure space of all systems considered, as well as its alterations, by using a dimensionality reduction procedure that makes it more intuitive to understand the high-dimensional space. Following ref. [31], we represent the local atom-centered environments of the structures through the smooth overlap of atomic positions (SOAP)[30] descriptors. We then obtain the similarity matrix between different conformers with the regularized entropy match kernel (REMatch).[29] We used SOAP descriptors with a cutoff of 5.0 Å, a Gaussian broadening of σ = 0.5 Å, and an intermediate regularization parameter γ = 0.01. SOAP kernels were calculated only for heavy atoms in the molecule (disregarding metal and hydrogen atoms) and were obtained using the GLOSIM package.[29,45]

For the dimensionality-reduced representation, here, we chose to use the metric multidimensional scaling (MDS) algorithm as implemented in the scikit-learn package.[46] This algorithm is similar to the Sketchmap algorithm previously used in ref. [31], but we found it more suitable for the data at hand, which comprise decorrelated local stationary points, instead of structures generated from molecular dynamics trajectories. In short, the low-dimensional map was obtained considering all calculated structures of Arg in the gas phase and through iterative minimization of the stress function

$$\sigma = \sum_{i \neq j} \left(D_{ij} - d_{ij}\right)^2, \tag{1}$$

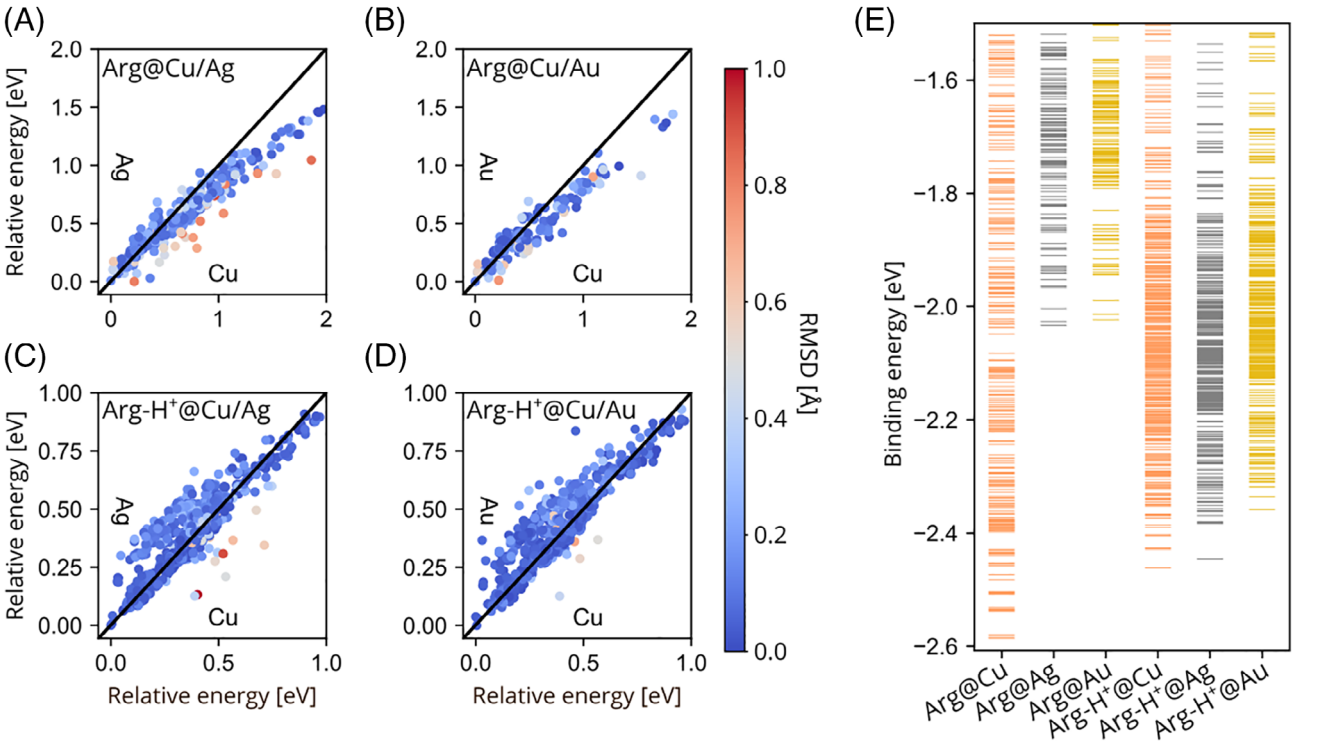


FIGURE 2 A-D, Correlation plots of relative energies of Arg or Arg-H^+ conformers on Cu, Ag, and Au (111) surfaces. Each dot corresponds to the same conformer optimized on the two surfaces addressed in each panel, color-coded with respect to the RMSD (heavy atoms only) between the superimposed optimized structures without taking surface atoms into consideration. E, Binding energies of Arg and Arg-H^+ on Cu(111), Ag(111), and Au(111) surfaces

where D_{ij} is the similarity between structures i and j in the high-dimensional space, and d_{ij} is the Cartesian distance in the low-dimensional (2D) space. We then projected structures in different environments onto the precomputed map of gas-phase Arg by fixing the parameters of the map and finding the low-dimensional coordinates x that minimize

$$\sigma_P(x) = \sum_{i=1}^{N_{Arg}} (|X - X_i| - |x - x_i|)^2 \quad (2)$$

where the sum runs over all structures in the optimized gas-phase Arg map, and X represents the high-dimensional representation of the new structure to be projected. In order to classify structural patterns, we use the following notations:

- We represent the protomers by the labels shown in Figure 1B and C.
- We identify the presence of strong intramolecular hydrogen bonds (H-bonds) whenever the distances between the hydrogen connecting donor and acceptor are below 2.5 Å. We label the H-bond pattern between two atoms in the molecules according to the nomenclature shown in Figure S3 in the SI.
- We further classify the structures according to the longest distance between two heavy atoms in the molecule.

The coordinates obtained as a result of the iterative metric MDS are not explicitly shown as axes on the plots as they are correlated to the descriptors used for the structural representation, which does not allow for a direct physical interpretation. These scatter plots just offer a visualization of the similarity matrix in lower dimensions.

2.3 | Binding energies and charge rearrangement

The binding energies for each structure were calculated as follows. For neutral Arg, we computed $E_b = E_{mol@surf} - E_{surf} - E_{mol}$, where $E_{mol@surf}$ corresponds to the total energy of the interface, E_{surf} is the total energy of the pristine metallic slab, and E_{mol} the total energy of the lowest-energy gas-phase conformer.

For charged Arg-H$^+$, we considered the binding energy of a two-step reaction. *First*, the interface is formed between the charged molecule and the clean surface: $E_{b1} = E_{mol}{}^{+}{}_{@surf} - E_{surf} - E_{mol}{}^{+}$, where $E_{mol}{}^{+}$ is the total energy of the most stable gas-phase conformer of the isolated charged molecule. *Second*, an electron from the metal neutralizes the unit cell where the adsorbed molecule is located, yielding $E_{b2} = E_{mol@surf} - E_{mol}{}^{+}{}_{@surf} - E_f$, where E_f corresponds to the Fermi energy of the metallic surface. The final binding energy is thus considered to be

$$E_b^+ = E_{b1} + E_{b2} = E_{mol@surf} - E_{surf} - E_f - E_{mol^+}. \quad (3)$$

For reference, we report the values we used for E_f at each surface in Table S5. These binding energies are shown in Figure 2D and will be discussed in detail in Section 3.

In addition, for conformers within 0.1 eV of the respective global minimum of each system, we have calculated free energies at finite temperatures within the harmonic approximation[47,48] in order to address the question about the thermal stability of such structures. We have calculated $F_{harm}(T) = E_{PES} + F_{vib}(T)$, where E_{PES} is the total energy obtained from DFT, and we have used textbook expressions for the harmonic vibrational Helmholtz free energy $F_{vib}(T)$. We have calculated the Hessian matrix, only taking into account displacements of the adsorbate. The surface was only considered an external field, which is a good approximation for physisorbed systems.

To address charge rearrangements after adsorption on the surface, we calculated electron density differences for selected structures on each surface by computing $\Delta\rho = \rho_{mol@surf} - \rho_{surf} - \rho_{mol}$ in the case of Arg and $\Delta\rho^{(+)} = \rho_{mol@surf} - \rho_{surf} - \rho_{mol}{}^{(+)}$, in the case of Arg-H$^+$. In these expressions, $\rho_{mol@surf}$ is the total electron density of the interface; ρ_{surf} is the electron density of the slab without molecule; and ρ_{mol} and $\rho_{mol}{}^{+}$ are electron densities of neutral Arg and charged Arg-H$^+$ molecules, respectively, with the same geometries as in the interface. The + sign denotes that the final density difference integrates to +1 electron in the case of Arg-H$^+$.

3 | RESULTS AND DISCUSSION

3.1 | The unconstrained structure space: Arg in isolation

We start by analyzing the unconstrained conformational space of Arg in isolation, which is formed by more than 1200 local stationary states.[21,44] In order to rationalize different structural arrangements in this space, we utilize the dimensionality-reduction algorithm discussed in Section 2.2 and

build a two-dimensional map. On this map, shown in Figure 3A, each dot represents one structure. A proximity of the dots implies similarities of the heavy-atom arrangement between the conformations. This is the low-dimensional map that is taken as a reference for comparison throughout this manuscript.

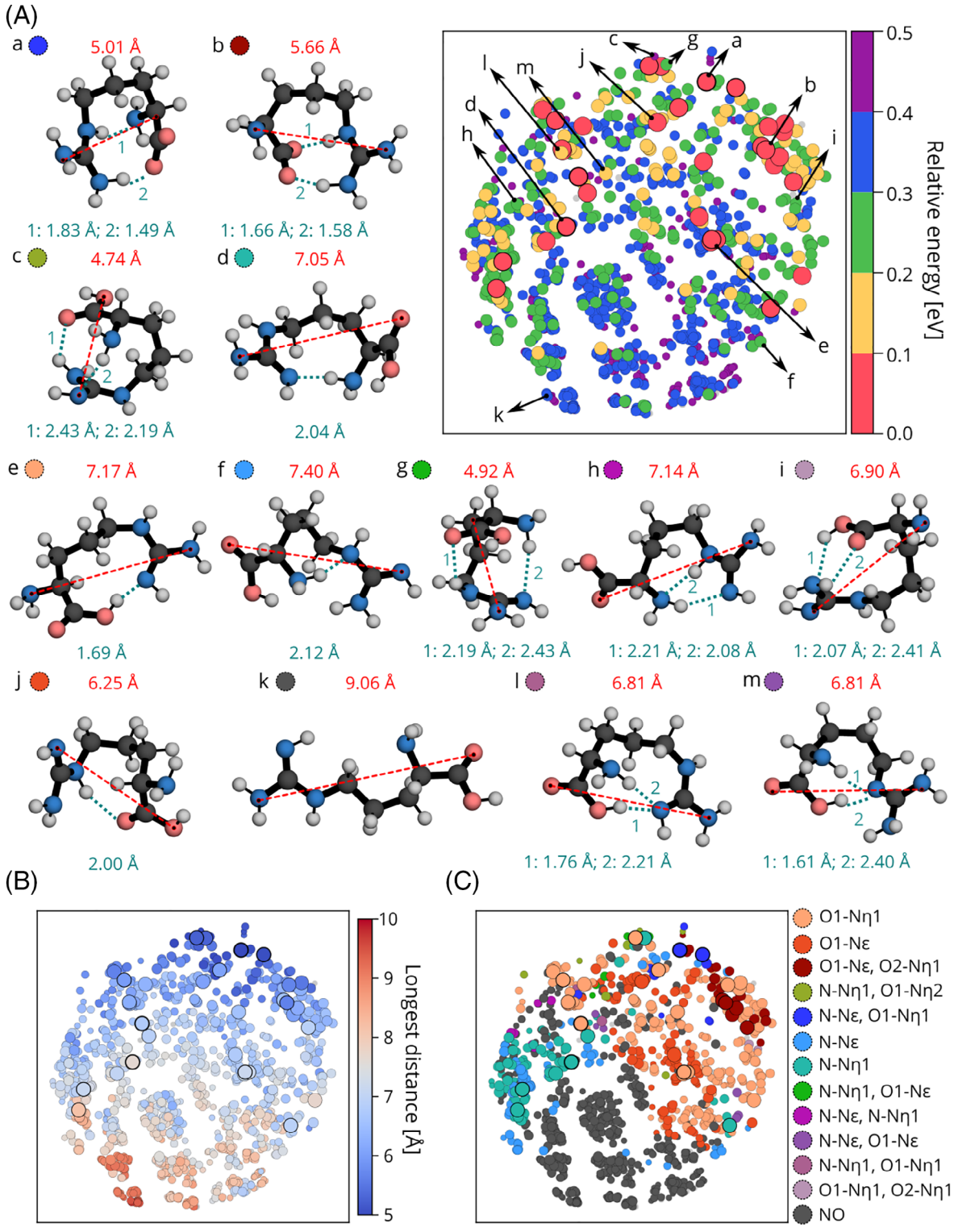


FIGURE 3 Low-dimensional map of Arg stationary points on the PES. Only points linked to structures with relative energy of 0.5 eV or lower are colored. Representative structures of all conformer families are visualized, as well as their H-bond distances (in turquoise) and the longest distance between two heavy atoms (in red) of the molecule. The maps are colored with respect to (A) relative energy, (B) longest distance, and (C) H-bond pattern. The size of the dots also reflect their relative energy, with larger dots corresponding to lower-energy structures

We proceed to color-code the dots on the map according to different properties. In Figure 3A, we show the map colored by the relative energy of each structure with respect to the global minimum ΔE_{rel}. We only color structures with ΔE_{rel} < 0.5 eV. The region with ΔE_{rel} < 0.1 eV is colored red and is represented by 32 different structures that occupy different parts of the map. The dominant protomer (98.6%) among these conformers is the one labeled **P1** in Figure 1, that is, nonzwitterionic. However, the lowest-energy structure, labeled *a* in panel (a) of Figure 3, is protomer **P3**, with a shared proton between the carboxylic and the guanidino group. This structure is compact, with the longest distance within the molecule of only 5.01 Å and presenting two strong intramolecular H-bonds. Zwitterionic protomers, denoted as **P4** and **P5** in Figure 1, do not appear in the gas phase.

Inspecting the map in Figure 3A, it is clear that low-energy conformers are almost exclusively present in the upper hemisphere of the plot. This can be rationalized in terms of the structural motifs that occupy these two halves of conformational space: In Figure 3B, we color-code the dots in terms of the longest extension of the conformers. While the upper hemisphere features compact structures, the lower hemisphere of the map is populated by extended conformers (with longest extensions between 7.5 Å and 10.0 Å). Many of them do not contain any H-bond or contain only one H-bond between the carboxyl and amino group. Extended conformers of Arg are energetically unfavored in the gas phase as the formation of strong H-bonds is crucial for the stabilization of Arg in isolation. Comparing the different plots in Figure 3, we see that all low-energy structures with ΔE_{rel} < 0.1 eV are indeed compact with one or two H-bonds.

In Figure 3C, we identify a total of 13 different families with respect to the number and character of H-bonds in the molecule, with ΔE_{rel} < 0.5 eV. Representative structures of all families are shown in panel (a). This family classification helps us understand why, in Figure 3A, there are structures of higher energies at similar regions as structures with lower energies. Even though these structures are typically in the same protomeric state and have a similar arrangement of the heavy atoms, the carboxyl group can rotate, giving rise to different H-bond patterns. These different arrangements can give rise to energy differences of up to 0.2 eV, as exemplified in Figure S4 in the SI. Including hydrogens in the SOAP descriptors used to build the 2D map could provide a better energy separation, but it would prevent us from comparing different protonation states, as shown in the next section.

3.2 | Adding a proton: Arg-H^+ in isolation

Arg-H^+ is of particular interest as it is the most abundant form of this amino acid under physiological pH conditions,[49] and we thus investigate changes of the conformational space introduced by the addition of a proton to the Arg amino acid. To that end, we plot a projection of all stationary points of the Arg-H^+ PES with ΔE_{rel} < 0.5 eV (referenced to its own global minimum) onto the map that was previously created for Arg. In Figure 4A, we color the dots in the map according to ΔE_{rel}, in Figure 4B according to the longest distance between heavy atoms in the molecule, and in Figure 4C according to the H-bond pattern. The gray dots in the maps represent all points in the Arg map of Figure 3 and are shown for ease of comparison.

The unique conformation types of Arg-H^+ can be grouped into eight different families in this energy range, which are represented in Figure 4A. Most families only have one H-bond, and there are no zwitterionic protomers. This means that, in isolation, only the protomer **P6** is populated. It is worth noting that, under physiological conditions (in solution), the zwitterionic protomer **P7** is preferred.

There are only two very similar structures (same family) with ΔE_{rel} < 0.1 eV in this case. The global minimum, labeled *a* in Figure 4A, contains two H-bonds within the molecule, between atoms N-Nε and O1-Nη (Figure 1). This particular structure resembles the lowest-energy structure of Arg with a proton added to the carboxyl group. This protonation results in an extension of the molecule by around 1 Å. That correlates with the location of the lowest-energy structure being slightly shifted on the map toward the region containing more extended structures.

We note that the structure space of Arg-H^+ is contained within the conformational space of Arg and also drastically reduced in numbers if compared to Arg: There are only 108 structures with ΔE_{rel} < 0.5 eV, compared to 1179 structures in the Arg case. In this energy range, regions of the map with very compact and very extended structures are not populated in this protonation state. This can be traced to the constraint imposed by the addition of the proton, which makes extended structures less stable due to the strong driving force to neutralize the charge imbalance created by the proton on the guanidino group. To rationalize why the most compact conformers are also less populated, we show, in Figure 5, the electron-density differences between the lowest-energy Arg-H^+ conformer and an Arg conformer created by fixing the same Arg-H^+ structure but neutralizing the charge and removing the hydrogen connected to the carboxyl group. This modification yields the same covalent connectivity observed in the global minimum of Arg. We show isosurfaces corresponding to electron accumulation in Arg-H^+ in red and electron depletion in Arg-H^+ (accumulation in Arg) in blue. We observe a density surplus between the O1 and Nη atoms in Arg, favoring the formation of a stronger H-bond leading to a more compact structure.

3.3 | Adsorption of Arg on Cu, Ag, Au (111) surfaces

We now turn to the analysis of the conformational space of Arg when in contact with metal surfaces, namely, Cu(111), Ag(111), and Au(111). In Figure 6, we show map projections of the stationary points with ΔE_{rel} < 0.5 eV (referenced to each respective global minimum) of Arg adsorbed

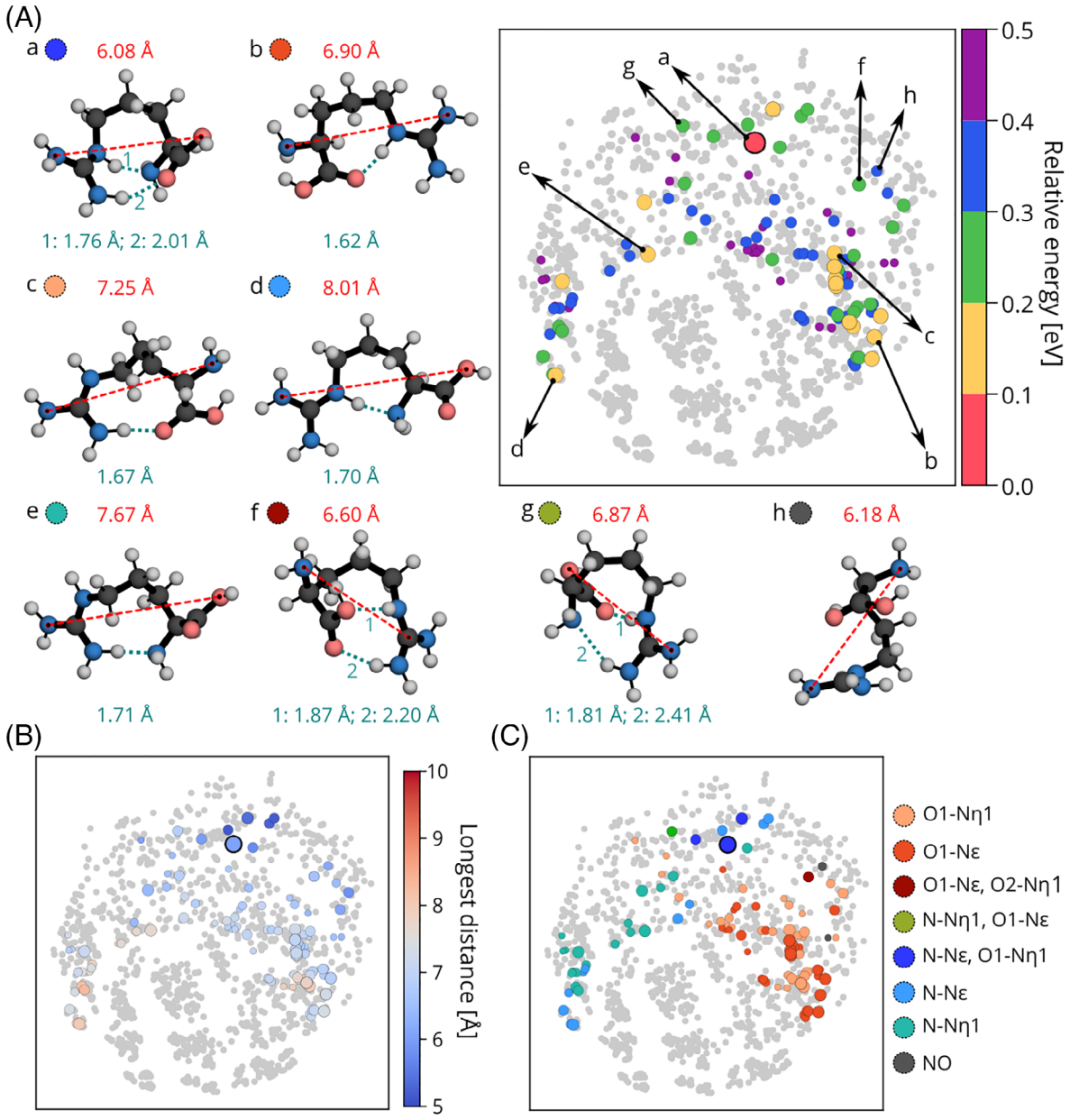


FIGURE 4 Representative conformers of the populated structure families within 0.5 eV of the global minimum of isolated Arg-H$^+$ and low-dimensional projections of all populated conformers onto the Arg map. Gray dots represent all structures from the original map of isolated Arg in Figure 3 and serve as a guide to the eye. The maps are colored with respect to (A) relative energy, (B) longest distance within the molecule, and (C) H-bond pattern

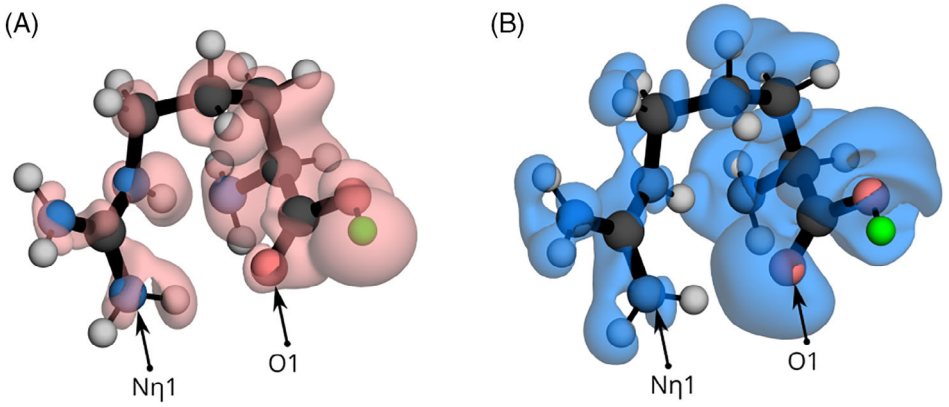


FIGURE 5 Electron density difference between Arg-H$^+$ and Arg calculated by neutralizing the charge and removing the hydrogen connected to the carboxyl group (marked in green) from the lowest-energy structure of Arg-H$^+$. The isosurfaces of electron density with value ±0.005 e/Bohr^3 corresponding to the (A) regions of electron accumulation on Arg-H$^+$ and (B) where the electron depletion on Arg-H$^+$, both compared to Arg

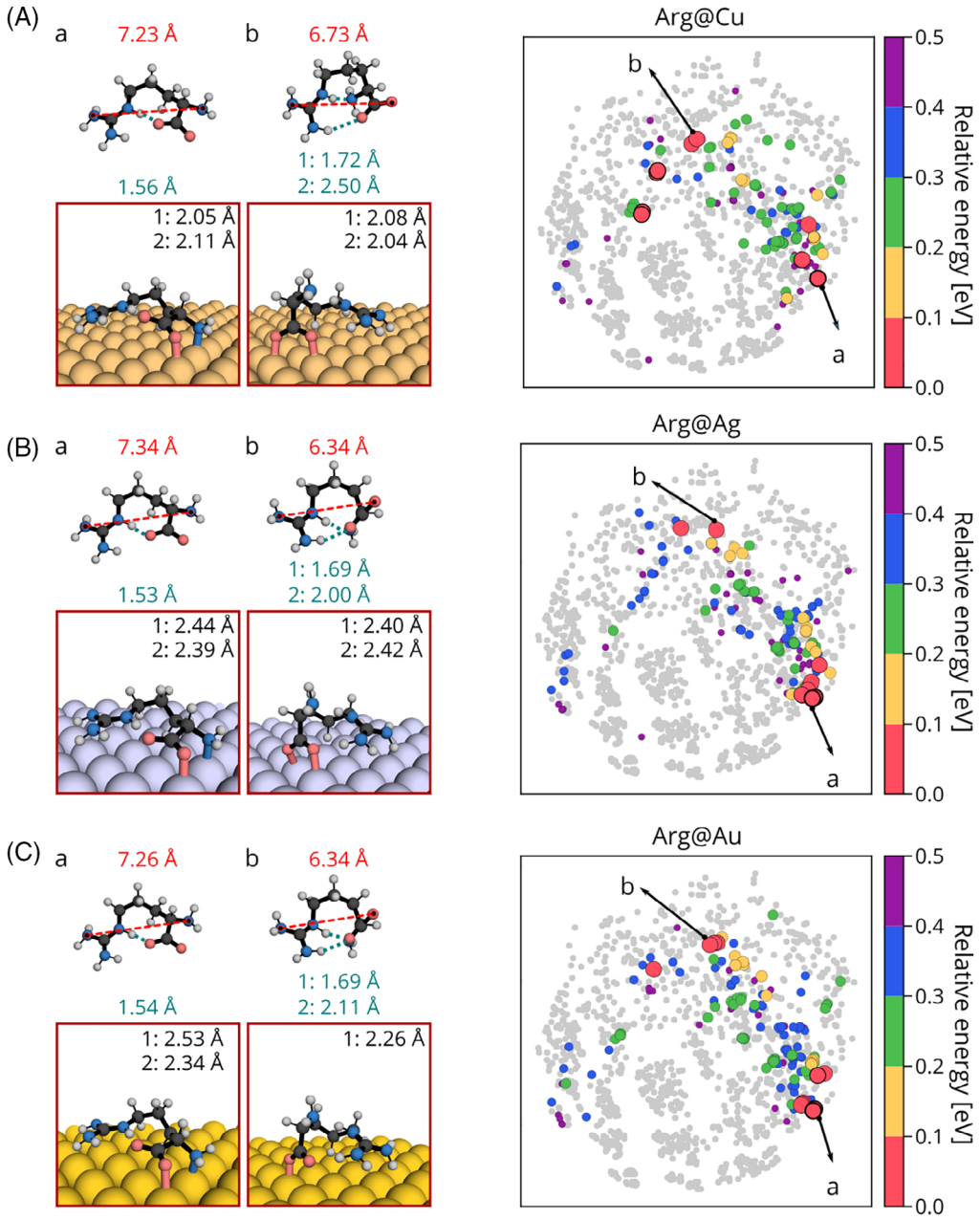


FIGURE 6 Low-dimensional projections of conformers of Arg adsorbed on (A) Cu(111), (B) Ag(111), and (C) Au(111) onto the gas-phase Arg map of Figure 3. Only conformers within 0.5 eV from the respective global minimum are colored. Gray dots represent all structures from the original map of gas-phase Arg and serve as a guide to the eye. In each panel, representative structures are shown from two perspectives: a side view where molecule and surface are shown (bottom) and the corresponding top view (top) where only the molecule is shown. The longest distance within each visualized conformer is reported in red, and H-bond lengths are reported in turquoise

on the three surfaces. The conformational space of Arg upon adsorption is reduced, and the conformers occupy similar regions of the map as the ones from Arg-H^+. We will learn in the following that this is mainly due to the formation of strong bonds with the surface that results in steric constraints of the space and also partially due to electron donation from the molecule to the metallic surfaces.

The lowest-energy structure lies on the same part of the map on all surfaces, which is different from the area where the gas-phase global minimum of Arg was located. These conformers, labeled *a* in Figure 6A-C, form a strong H-bond between atoms O1 and Nε. The longest distance within the molecule lies between 7.20 and 7.35 Å. This structure binds most strongly to all three surfaces through its amino and carboxyl groups.

Other low-energy structures in all surfaces form strong bonds to the surfaces only through the carboxyl group, as exemplified by the structure labeled *b* in all panels of Figure 6. These bonds are formed most favorably on *top* positions, that is, vertically on top of a surface metal atom. In particular, for Cu(111), the atomic spacing of the Cu atoms on the surface favors both oxygens to bind on *top* positions simultaneously.

The favorable formation of these bonds is connected to the fact that all conformers with $\Delta E_{rel} < 0.2$ eV are in the protomeric state **P3**, in which the carboxyl group is deprotonated. The bonds with the surface and a favorable vdW attraction effectively flatten the molecular conformation, thus energetically favoring more elongated structures. Protomers of type **P1**, which were dominant in the gas phase, only appear with $\Delta E_{rel} > 0.3$ eV on Cu and Ag and with $\Delta E_{rel} > 0.2$ eV on Au. Protomers **P4** and **P5** are again not observed. Regarding the intramolecular H-bond patterns, within 0.5 eV from the global minimum, we can identify seven different families on Cu(111) and six families on both Ag(111) and Au(111). These families contain H-bonds where the carboxyl group predominantly participates. All families are represented in Table S4 in the SI.

3.4 | Adsorption of Arg-H$^+$ on Cu, Ag, Au (111) surfaces

Finally, we characterize the conformational-space changes arising from the simultaneous addition of a proton and the adsorption onto metallic surfaces. In Figure 7, we show the projection of Arg-H$^+$ conformers adsorbed on Cu, Ag, and Au(111) onto the map of isolated Arg. These projections, particularly the comparison of the plots in Figures 6 and 7, reveal that the conformational space of adsorbed Arg-H$^+$ is larger than the one of adsorbed Arg. While Arg-H$^+$ features more than 500 conformers within $\Delta E_{rel} < 0.5$ eV, Arg only counts about 150 conformers in the same energy range. Interestingly, the adsorption of Arg-H$^+$ to a metal surface also results in an increase of the occupied structure space in comparison to isolated Arg-H$^+$ (108 structures with $\Delta E_{rel} < 0.5$ eV), shown in Figure 4. In fact, the structures occupy similar regions of the map as the ones occupied by Arg-H$^+$, with the addition of extended structures that are located at the bottom of the map.

We identify four different families on Cu(111) and three on Ag(111) and Au(111) with $\Delta E_{rel} < 0.1$ eV. Representative conformers of these families are shown in Figure 7. The lowest-energy conformer, labeled *a* in Figure 7A-C, appears on all surfaces in the same region of the map as for adsorbed Arg. The largest distance within the molecule lies around 7 Å, and it also has a strong H-bond linking the carboxyl-O and the Nε atoms. The structure, however, does not present the same orientation to the surface compared to the lowest-energy conformer of Arg and does not form strong bonds with the surface. With the exception of the extended structure on Cu(111), labeled *d* in Figure 7A, all conformers with $\Delta E_{rel} < 0.1$ eV on all surfaces contain one intramolecular H-bond involving either carboxyl-O and Nε atom (labeled a), backbone N and Nε atoms (labeled b), or carboxyl O and an Nη atom (labeled c). Compared to adsorbed Arg, adsorbed Arg-H$^+$ structures become, on average, 1.0 Å more extended (see SI). The guanidino and carboxyl groups often lie parallel to the surface. The protomer **P6**, the only one present in the gas phase, is also dominantly populated on the surfaces. However, we do observe very few conformers in the zwitterionic **P7** state. These structures are at least 0.2 eV higher in energy than the global minimum.

With respect to the number of bonds that Arg-H$^+$ forms with the surface, the picture is very different from adsorbed Arg. Within the lower 0.15 eV, we do not observe short (strong) bonds of O or N atoms to the surfaces. This lack of constraint by the surface contributes to the increased structure space of adsorbed Arg-H$^+$. In addition, the molecule accepts electrons from the surface, becoming less positively charged, as we discuss in detail in the next section. We conclude that Arg-H$^+$ interacts with the metallic surfaces mostly through van der Waalsand electrostatic interactions.

3.5 | Electronic structure and trends across surfaces

In the previous sections, we focused on structural aspects of the adsorbed molecules and the most prominent bonds the molecules make with the metallic surfaces. In the following, we will discuss different aspects of the molecule-surface interactions with the goal of identifying trends across these systems.

The binding energies for all surfaces were calculated as discussed in Section 2.3. The larger negative values in Figure 2D correspond to the stronger binding of the molecule to the surface. In the case of adsorbed Arg, many conformers bind to Cu more strongly than to Ag and Au. As mentioned previously, Arg forms strong bonds to the surfaces, but the binding of the deprotonated carboxyl group of Arg to the Cu(111) surface is geometrically favored as discussed above. In the case of adsorbed Arg-H$^+$, there is no pronounced difference in binding strengths to the surfaces, and the values are comparable to binding energies obtained for Arg adsorbed on Cu(111). This correlates with the observation that the interaction of Arg-H$^+$ to the surfaces happens mostly through dispersion and electrostatic interactions. Despite the strong binding to the surface, it is also seen in Figure 2A that the interaction of Arg-H$^+$ with the surface does not strongly template the conformations of this molecule, implying a low corrugation (that is, homogeneity) of the molecule-surface interaction and allowing for a larger variety of conformers with similar energies. This is in contrast to the molecule-surface interaction of Arg, which is more inhomogeneous due to the formation of bonds through specific chemical groups. In addition, we have estimated harmonic vibrational free energies for all conformers with $\Delta E_{rel} < 0.1$ eV on each surface. In contrast to what has been reported for longer helical peptides,[22,24] the global minimum remains the same in all cases, as reported in Figure S6 in the SI. For Arg-H$^+$, we observe relative energy rearrangements of up to 50 meV at 300 K, which changes the relative energy hierarchy of conformers less stable than the global minimum.

We then focus on the distance between the molecule and the surfaces. We define this quantity by measuring the distance of the center of mass (COM) of the molecule with respect to the surface plane defined by the top layer of surface atoms. These distances are collected

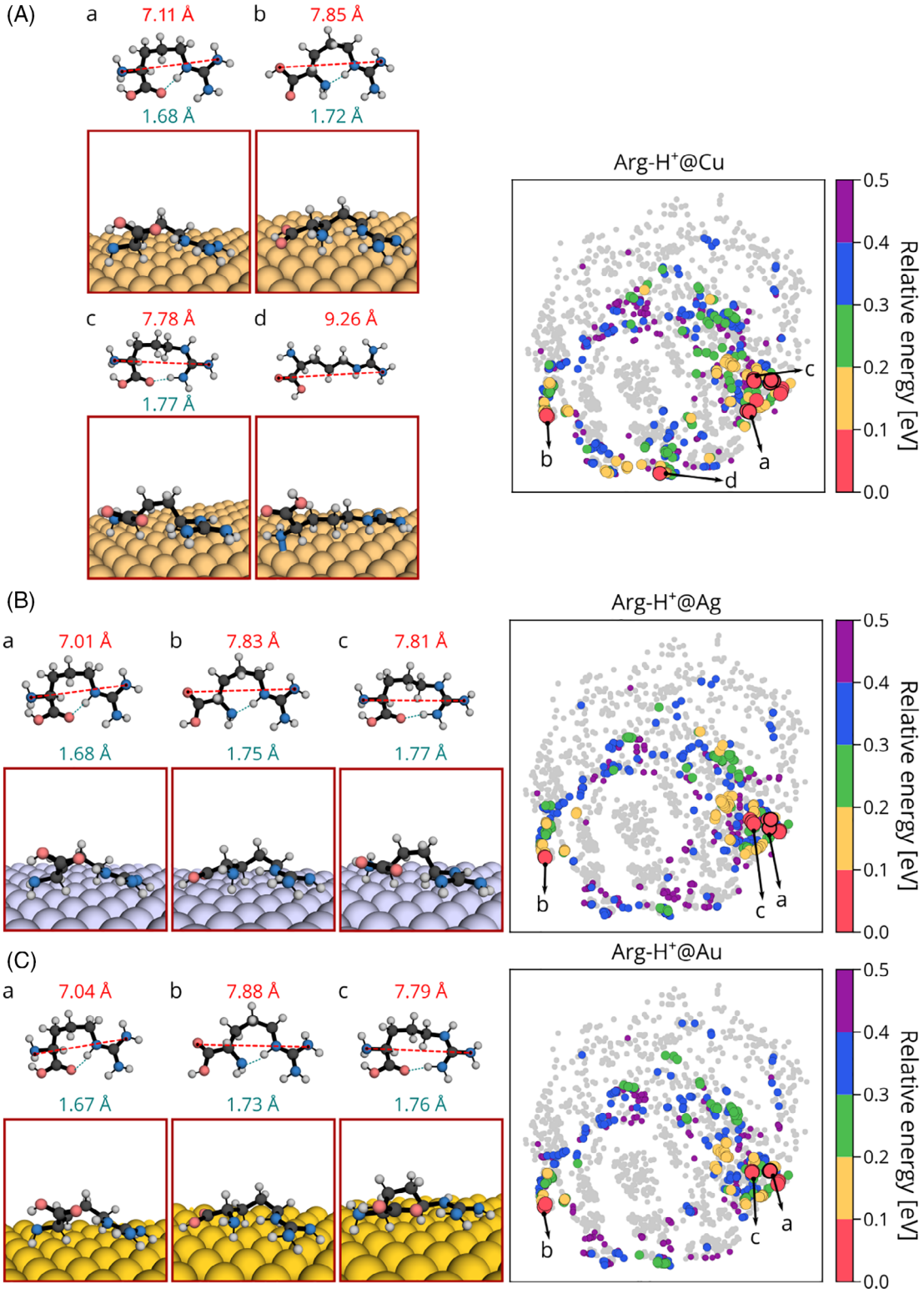


FIGURE 7 Low-dimensional projections of conformers of Arg-H^+ adsorbed on (A) Cu(111), (B) Ag(111), and (C) Au(111) onto the gas-phase Arg map of Figure 3. Only conformers within 0.5 eV from the respective global minimum are colored. Gray dots represent all structures from the original map of gas-phase Arg and serve as a guide to the eye. In each panel, representative structures are shown from two perspectives: a side view where molecule and surface are shown (bottom) and the corresponding top view (top) where only the molecule is shown. The longest distance within each visualized conformer is reported in red, and H-bond lengths are reported in turquoise

in Figure 8. The COM is closer to Cu(111) than to Ag(111) and Au(111) for both Arg and Arg-H$^+$. In addition, on all surfaces, Arg lies closer than Arg-H$^+$, in agreement with the observation that Arg forms stronger bonds to the surface. The extended structures of Arg-H$^+$, at the bottom of the maps, tend to be closer to the surface than those that have H-bonds within the molecule, likely due to the stronger van der Waals (vdW) attraction to the surface by extended conformations.

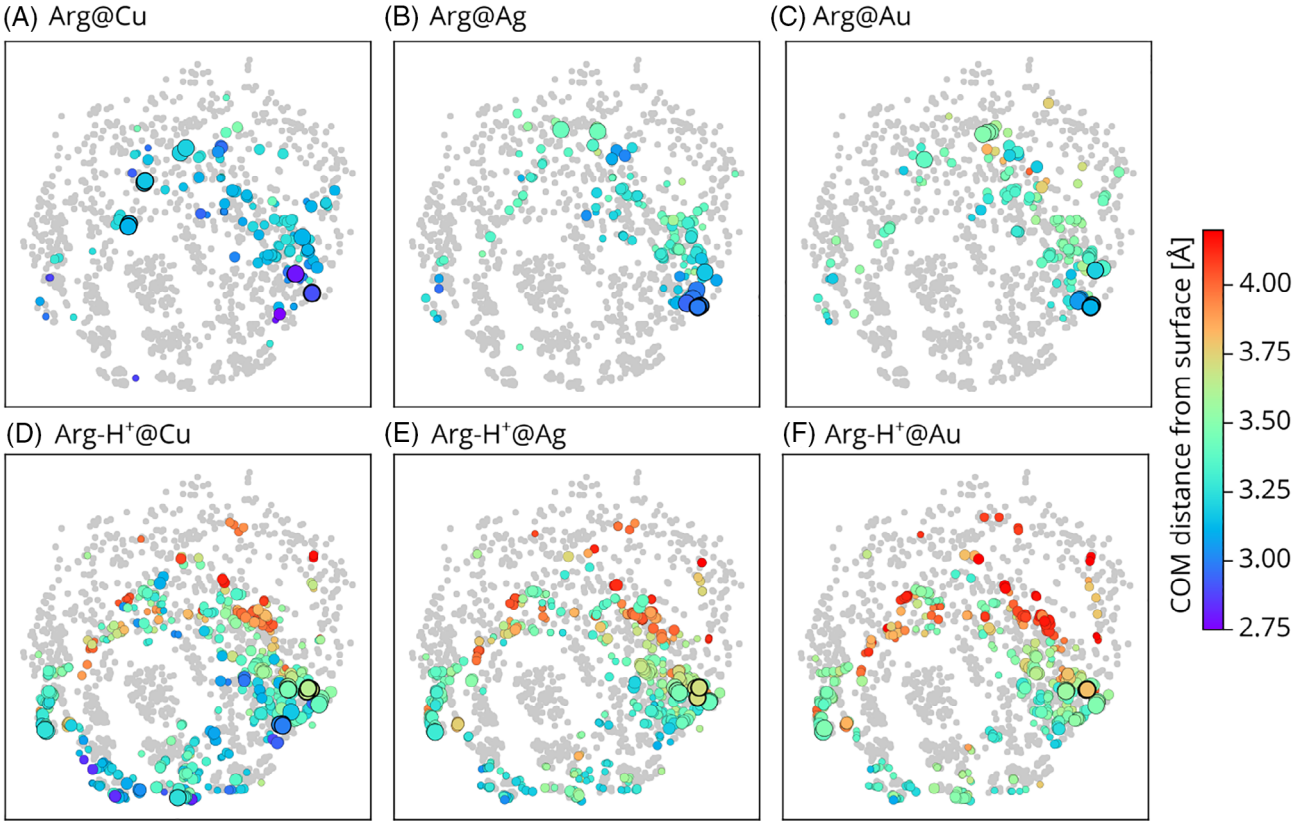
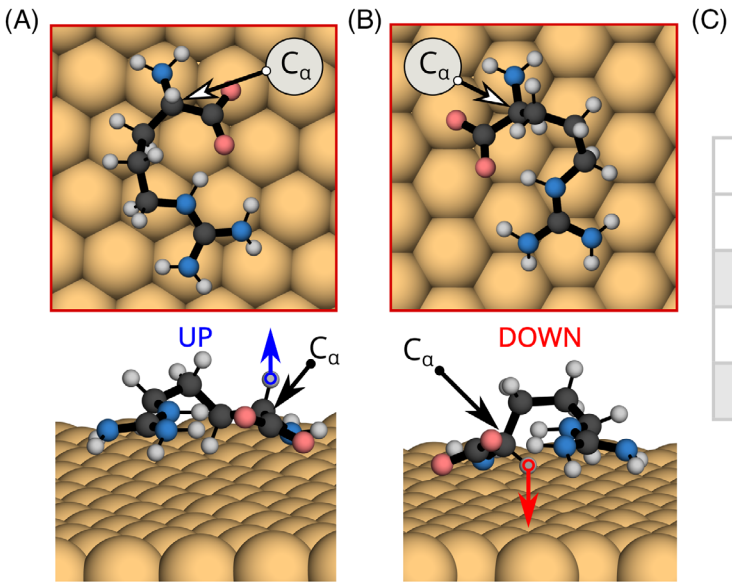


FIGURE 8 Low-dimensional projections of adsorbed Arg and Arg-H$^+$ on Cu(111), Ag(111), and Au(111) color-coded with respect to the distance from the center of mass of the molecule to the surface. Gray dots represent all structures from the original map of isolated Arg where the projection was made and serve as a guide to the eye



	Arg		Arg-H$^+$	
	up	*down*	*up*	*down*
Cu	12/100	0/48	8/255	13/355
Ag	10/110	0/50	3/255	10/292
Au	13/104	1/53	2/219	12/284

FIGURE 9 Orientation of the $C_\alpha H$ group in (A) *up* orientation (hydrogen pointing toward vacuum) and (B) *down* orientation (hydrogen pointing toward the surfaces). (C) The number of structures with *up* and *down* orientation within 0.1/0.5 eV from the global minimum of each surface

The difference in COM distances to the surfaces between Arg and Arg-H^+ is apparently related to the preferred orientation of the chiral center of the molecule to the surface. The chiral C_α carbon can point its bonded hydrogen toward the surface (labeled *down* in the following) or toward the vacuum region (labeled *up* in the following). Examples of different molecular orientations are shown in Figure 9A. The dominant orientation with respect to the surface is different in the cases of Arg and Arg-H^+, as evidenced by the numbers presented in Figure 9B. The lower-energy structures are mostly in the *up* orientation for Arg and mostly in the *down* orientation for Arg-H^+ (see also map in Figure S7). As discussed in the previous sections, despite showing different orientations of their $C_\alpha H$ groups, the lowest-energy structures for both molecules adsorbed on each surface have very similar conformations. As the addition or removal of a proton can apparently alter the preference of the chiral-center orientation, we propose that it could template different chiralities of self-assembled superstructures on the surface.[50]

We then investigated the rearrangement of the electronic density upon binding of the molecules to the different surfaces. In Figure 10, we show the electronic density rearrangement created by the lowest-energy conformer at each surface, integrated over the axis parallel to the surface, overlaid on the side view of the 3D density rearrangement. In addition, we show a top view of the density rearrangement in each case. Examples of further conformers are summarized in the SI, Figures S8-S13. The data show that Arg donates electrons to the surface, while Arg-H^+ accepts electrons from the surface. We have checked this propensity for selected conformers by the integration of the electronic density

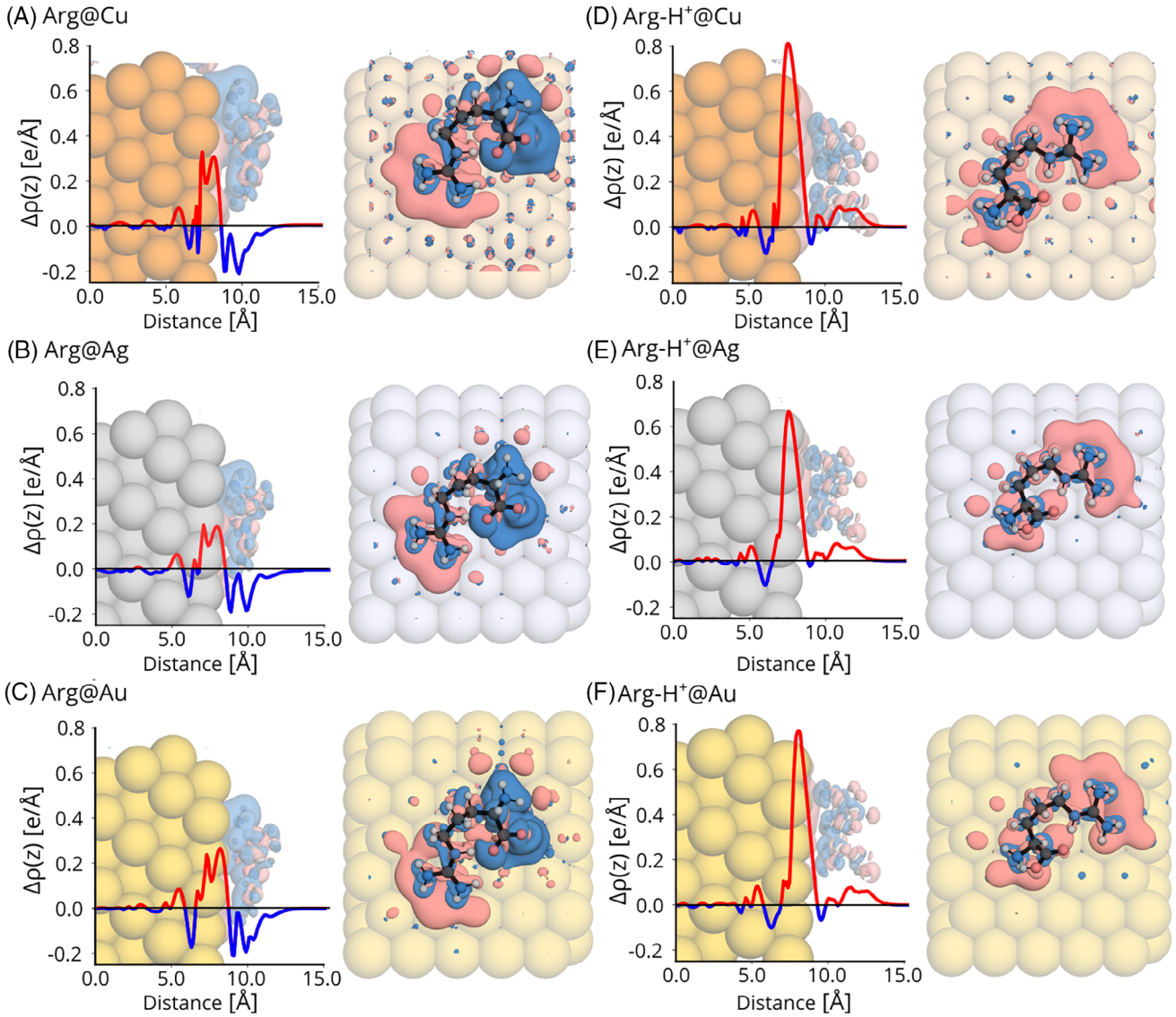


FIGURE 10 Electronic-density difference averaged over the directions parallel to the surface for the lowest-energy conformers of Arg adsorbed on Cu(111) (A), Ag(111) (B), and Au(111) (C), as well as of Arg-H^+ adsorbed on Cu(111) (D), Ag(111) (E), and Au(111) (F). Positive values (red) correspond to electron density accumulation, and negative values (blue) correspond to electron density depletion. In each panel, we also show side and top views of the 3D electronic density rearrangement. Blue isosurfaces correspond to an electron density of +0.05 e/Bohr^3 and red isosurfaces to −0.05 e/Bohr^3

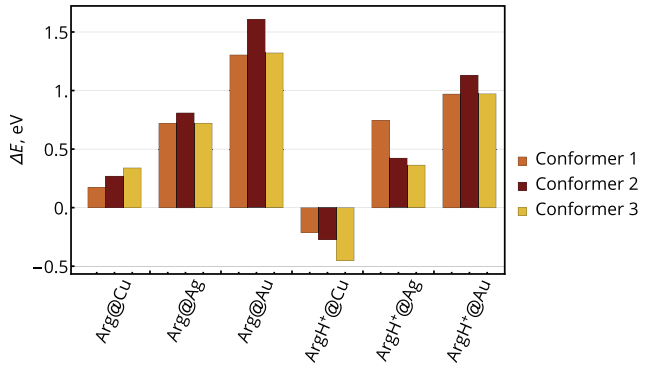


FIGURE 11 Energy differences upon hydrogen dissociation for selected conformers (see text and SI) of Arg and Arg-H^+ on all metallic surfaces. $\Delta E = E_{dep} - E$, where E_{dep} is the total energy of the dissociated structure after optimization (including the adsorbed hydrogen), and E is the energy of the optimized intact structure. A negative ΔE indicates that deprotonation is favored

rearrangement around the molecule and by calculating the Hirshfeld charge remaining on the molecule for the full database (see SI, Table S6). When comparing Hirshfeld charges on the molecule and those obtained from the electronic density rearrangement, we observe that Hirshfeld charges are always underestimated by 0.3 to0.5 e. In addition, we observe that the depletion and accumulation of charge is not uniform through the lateral extension of the molecule. This behavior is consistent with the level alignment predicted by the PBE Kohn-Sham energy levels, as shown in Figure S14 in the SI. However, we note that quantitative values of charge transfer are often inaccurate at this level of theory, as characterized in refs. [51,52]. Optimally tuned range-separated hybrid functionals would yield more accurate values, but their computational cost is prohibitive for use in this whole database. Nevertheless, hybrid-functional calculations of selected conformers (see SI, Figure S15) confirm the qualitative trend. Therefore, we conclude that the protonation state again critically impacts these systems, in this case by qualitatively changing the redistribution of electronic charge.

It was observed experimentally that amino acids can undergo deprotonation on reactive surfaces.[53–58] Here, we also investigated whether the deprotonation of Arg and Arg-H^+ was favorable on any of the surfaces studied here; the details of the procedure can be found in the SI. The results are summarized in Figure 11. They show that deprotonation of Arg-H^+ is favorable in Cu(111), such that Arg-H^+ would be predominantly deprotonated. The population of deprotonated Arg on Cu(111) could reach about 0.1% of the molecules at room temperature from a simple Arrhenius estimation. Given that we have not observed any spontaneous dissociation upon optimization of Arg-H^+ on Cu(111), we conclude that, although favorable, this dissociation of H does not occur without a barrier. In all other surfaces, the barrier for dissociation would be rather high for both molecules.

4 | CONCLUSIONS

In this paper, we have characterized the conformational space of the arginine amino acid in its neutral and protonated form in different non-biological environments, that is, in isolation and in contact with metallic surfaces. In particular, we have analyzed how and why different parts of the conformational space become accessible or excluded depending on the protonation state and the environment, showing the importance of bond formation and charge rearrangement in these systems.

This study included the construction of a database based on thousands of structures optimized by DFT, including dispersion interactions. The construction of this database is a result in itself. We found, for example, that it is advantageous to start from a comprehensive sampling of the conformational space of the least-constrained molecular form, which in our case was the neutral Arg amino acid in the gas phase. This is evidenced by the fact that, in our low-dimensional projections, all low-energy conformers we observe on the surfaces, for both Arg and Arg-H^+, lie among structural conformations that were already present in the gas-phase sampling of Arg, albeit often with high relative energies. In addition, we find that, within Cu, Ag, and Au surfaces, relative energies between different conformers are largely preserved when changing the substrate.

In recent years, much effort has been invested in developing and improving classical force fields (FF) for simulation of organic-inorganic interfaces, including peptides on surfaces.[59–61] However, as we discuss in the SI, the existing force fields are not accurate for these scenarios where amino acids are in contact with metal surfaces and under vacuum conditions. Nevertheless, these situations are extremely relevant for STM experiments[11–13,62] and other technological applications.[1,2] As we illustrate in the SI, even though these force fields can sample the relevant areas of conformational space, they are not able to capture consistent energy hierarchies. In addition, the molecular chemical groups show a preference for adsorbing on different surface sites, which could have a considerable impact on self-assembly studies. Databases such as this one will serve as an important source of data for further parametrization and improvement of these potentials.

Regarding the structural space of Arg and Arg-H^+ adsorbed on (111) surfaces of Cu, Ag, and Au, we have learned the following: The adsorption of Arg leads to the formation of strong bonds on the surface that involve mostly the carboxyl and amino groups. This stabilizes the protomer that we label **P3** in this work, where the carboxyl group is deprotonated, and the side chain is protonated. This is different from the dominant

protomer in the gas phase **P1**. The bonds on the surface sterically constrain the conformations of this molecule, thus decreasing the number of structures with respect to the numbers observed in the gas phase. The molecule also donates electrons to the surface, becoming slightly positively charged. We do not observe fully extended structures lying on the surface, and most conformers exhibit intramolecular H-bonds. The majority of conformers of Arg in the low-energy region adsorbs with the $C_\alpha H$ chiral center pointing the hydrogen atom away from the surfaces.

Arginine in its protonated form, that is, Arg-H^+, is the most abundant form of this amino acid in biological environments, where it typically adopts the zwitterionic protomer **P7**. In the gas phase, we observe that the nonzwitterionic state **P6** is dominant and that the addition of the proton decreases the number of allowed conformations with respect to isolated Arg due to the added electrostatic interactions and the passivation of the carboxyl group that would otherwise be involved in intramolecular H-bonds. Upon adsorption on the metallic surfaces, we observe that the protomer **P6** is still dominant and that there are no strong bonds formed on the surface. In addition, this molecule receives electrons from the surface, thus becoming less positively charged. Both effects conspire to yield a homogeneous (flat) molecule-surface interaction and a relatively high population of different structures in the low-energy range. Contrary to Arg, most low-energy conformers of Arg-H^+ adsorb with the C_α-H chiral center pointing the hydrogen atom towards to the surfaces. Finally, through the calculation of dissociation energies, we also conclude that the deprotonation of Arg-H^+ is energetically favorable only on Cu(111).

Our observations regarding the preferred protomers and deprotonation propensities discussed above are consistent with the observations in the literature that the adsorption of amino acids in their anionic and deprotonated form is common on reactive metals such as Cu(111).[8] One pronounced difference that we find among surfaces is the average adsorption height of the molecules: They follow the trend Cu(111) < Ag(111) < Au(111), and Arg is always closer than Arg-H^+ to the same respective surface.

The set of electronic-structure calculations presented here shows that a flexible amino acid such as Arginine presents a rich conformational space involving different protomeric states and molecule orientations with respect to the surface, allied to a complex charge rearrangement. Going forward, it is clear that the likes of this study based solely on DFT cannot become a routine method due to the elevated computational cost. Addressing the whole breadth of amino acids, as well as self-assembly of these structures on surfaces, will profit from using this study as a benchmark and a means to develop models, possibly based on different machine-learning techniques,[63–69] that can bypass the cost of thousands of DFT structure optimizations.

ACKNOWLEDGEMENTS

This work was supported by the Max Planck-EPFL Center for Molecular Nanoscience and Technology. We acknowledge Michele Ceriotti for insightful discussions and critical comments on the manuscript. We also acknowledge fruitful discussions with Yair Litman, Stephan Rauschenbach, and Sabine Abb. M. R. also acknowledges support from the Max Planck Research Network on Big-Data-Driven Science (BigMAX).

AUTHOR CONTRIBUTIONS

Dmitrii Maksimov: data curation; formal analysis; investigation; methodology; validation; visualization; writing-original draft; writing-review and editing. Carsten Baldauf: conceptualization; formal analysis; investigation; methodology; supervision; writing-original draft; writing-review and editing. Mariana Rossi: conceptualization; formal analysis; funding acquisition; methodology; resources; supervision; validation; writing-original draft; writing-review and editing.

DATA ACCESSIBILITY

The data presented here have been uploaded to the NOMAD repository[32] and can be visualized in Chemiscope[70] through the interface hosted in https://th.fhi-berlin.mpg.de/projects/chemiscope/.

ORCID

Dmitrii Maksimov https://orcid.org/0000-0003-4448-8848
Carsten Baldauf https://orcid.org/0000-0003-2637-6009
Mariana Rossi https://orcid.org/0000-0002-3552-0677

REFERENCES

[1] Y. Wang, M. Lingenfelder, S. Fabris, G. Fratesi, R. Ferrando, T. Classen, K. Kern, G. Costantini, *J. Phys. Chem. C* **2013**, *117*(7), 3440.
[2] C. Guo, X. Yu, S. Refaely-Abramson, L. Sepunaru, T. Bendikov, I. Pecht, L. Kronik, A. Vilan, M. Sheves, D. Cahen, *Proc. Nat. Acad. Sci.* **2016**, *113*(39), 10785.
[3] D. Khatayevich, T. Page, C. Gresswell, Y. Hayamizu, W. Grady, M. Sarikaya, *Small* **2014**, *10*(8), 1505.
[4] M. S. Mannoor, H. Tao, J. D. Clayton, A. Sengupta, D. L. Kaplan, R. R. Naik, N. Verma, F. G. Omenetto, M. C. McAlpine, *Nat. Commun.* **2012**, *3*, 763.
[5] O. J. Guy, G. Burwell, Z. Tehrani, A. Castaing, K. A. Walker, S. H. Doak, *Mater. Sci. Forum* **2012**, *711*, 246.
[6] X. Zhao, F. Pan, J. R. Lu, *Prog. Nat. Sci.* **2008**, *18*(6), 653.
[7] M. Sarikaya, C. Tamerler, A. K. Y. Jen, K. Schulten, F. Baneyx, *Nat. Mater.* **2003**, *2*(9), 577.
[8] D. Costa, C. M. Pradier, F. Tielens, L. Savio, *Surface Sci. Rep.* **2015**, *70*(4), 449.

[9] H. Heinz, H. Ramezani-Dakhel, *Chem. Soc. Rev.* **2016**, *45*(2), 412.
[10] T. R. Walsh, M. R. Knecht, *Chem. Rev.* **2017**, *117*(20), 12641.
[11] S. Rauschenbach, G. Rinke, R. Gutzler, et al., *ACS Nano* **2017**, *11*(3), 2420.
[12] S. Rauschenbach, M. Ternes, L. Harnau, K. Kern, *Ann. Rev. Anal. Chem.* **2016**, *9*(1), 473.
[13] S. Abb, L. Harnau, R. Gutzler, S. Rauschenbach, K. Kern, *Nat. Commun.* **2016**, *7*, 10335.
[14] C. Méthivier, H. Cruguel, D. Costa, C. M. Pradier, V. Humblot, *Langmuir* **2016**, *32*(51), 13759.
[15] R. Di Felice, A. Selloni, E. Molinari, *J. Phys. Chem. B* **2003**, *107*(5), 1151.
[16] R. Di Felice, A. Selloni, *J. Chem. Phys.* **2004**, *120*(10), 4906.
[17] L. M. Ghiringhelli, P. Schravendijk, L. Delle Site, *Phys. Rev. B* **2006**, *74*(3), 35437.
[18] C. Arrouvel, B. Diawara, D. Costa, P. Marcus, *J. Phys. Chem. C* **2007**, *111*(49), 18164.
[19] F. Iori, S. Corni, R. Di Felice, *J. Phys. Chem. C* **2008**, *112*(35), 13540.
[20] G. Hong, H. Heinz, R. R. Naik, B. L. Farmer, R. Pachter, *ACS Appl. Mater. Inter.* **2009**, *1*(2), 388.
[21] M. Ropo, V. Blum, C. Baldauf, *Sci. Rep.* **2016**, *6*, 35772.
[22] M. Rossi, M. Scheffler, V. Blum, *J. Phys. Chem. B* **2013**, *117*(18), 5574.
[23] M. Rossi, S. Chutia, M. Scheffler, V. Blum, *J. Phys. Chem. A* **2014**, *118*(35), 7349.
[24] F. Schubert, M. Rossi, C. Baldauf, K. Pagel, S. Warnke, G. von Helden, F. Filsinger, P. Kupser, G. Meijer, M. Salwiczek, B. Koksch, M. Scheffler, V. Blum, *Phys. Chem. Chem. Phys.* **2015**, *17*(11), 7373.
[25] C. Baldauf, M. Rossi, *J. Phys.: Cond. Matter* **2015**, *27*(49), 493002.
[26] M. Ceriotti, G. A. Tribello, M. Parrinello, *Proc. Nat. Acad. Sci.* **2011**, *108*(32), 13023.
[27] G. A. Tribello, M. Ceriotti, M. Parrinello, *Proc. Nat. Acad. Sci. U S A* **2012**, *109*(14), 5196.
[28] M. Ceriotti, G. A. Tribello, M. Parrinello, *J. Chem. Theory. Comput.* **2013**, *9*(3), 1521.
[29] S. De, A. P. Bartók, G. Csányi, M. Ceriotti, *Phys Chem Chem Phys* **2016**, *18*, 13754.
[30] A. P. Bartók, S. De, C. Poelking, et al., *Sci. Adv.* **2017**, *3*(12), e1701816.
[31] S. De, F. Musil, T. Ingram, C. Baldauf, M. Ceriotti, *J. Cheminformat.* **2017**, *9*(1), 6.
[32] Maksimov D, Baldauf C, Rossi M, *Database of Arg and Arg-H$^+$ adsorbed on Cu(111), Ag(111) and Au(111) in the NOMAD repository*, **2020**. https://dx.doi.org/10.17172/NOMAD/2020.03.24-1.
[33] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, M. Scheffler, *Comput. Phys. Commun.* **2009**, *180*, 2175.
[34] V. Havu, V. Blum, P. Havu, M. Scheffler, *J. Comput. Phys.* **2009**, *228*(22), 8367.
[35] Hermann J, Tkatchenko A, Density functional model for van der Waals interactions: Unifying many-body atomic approaches with nonlocal functionals. *Phys. Rev. Lett.* arXiv e-prints **2020**, *124*, 146401. https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.124.146401.
[36] V. G. Ruiz, W. Liu, A. Tkatchenko, *Phys. Rev. B* **2016**, *93*(3), 035118.
[37] V. G. Ruiz Lopez. Adsorption of atoms and molecules on surfaces: density-functional theory with screened van der Waals interactions. *Doctoral thesis*, Technische Universitaet Berlin. **2016**.
[38] W. Liu, V. G. Ruiz, G. X. Zhang, B. Santra, X. Ren, M. Scheffler, A. Tkatchenko, *N. J. Phys.* **2013**, *15*(5), 53046.
[39] W. Liu, J. Carrasco, B. Santra, A. Michaelides, M. Scheffler, A. Tkatchenko, *Phys Rev B* **2012**, *86*(24), 245405.
[40] W. A. Al-Saidi, H. Feng, K. A. Fichthorn, *Nano Lett.* **2012**, *12*(2), 997.
[41] J. van Ruitenbeek, *Nat. Mater.* **2012**, *11*, 834.
[42] C. Wagner, N. Fournier, F. S. Tautz, R. Temirov, *Phys Rev Lett* **2012**, *109*, 076102.
[43] J. Carrasco, W. Liu, A. Michaelides, A. Tkatchenko, *J. Chem. Phys.* **2014**, *140*(8), 84704.
[44] M. Ropo, M. Schneider, C. Baldauf, V. Blum, *Sci. Data* **2016**, *3*, 160009.
[45] M. Ceriotti, S. De, F. Musil, Glosim package; https://github.com/cosmo-epfl/glosim (accessed: April 2020).
[46] F. Pedregosa, G. Varoquaux, A. Gramfort, et al., *J. Machine Learn. Res.* **2011**, *12*, 2825.
[47] D. A. McQuarrie, *Statistical Mechanics*, Calif University Science Books, Sausalito **2000**.
[48] B. Fultz, *Prog. Mater. Sci.* **2010**, *55*(4), 247.
[49] M. B. Bolger, *Introduction to biophysical methods for protein and nucleic acid research*, (Eds: J. A. Glasel, M. P. Deutscher) Academic Press, San Diego **1995**, Chapter 9–Computational Techniques in Macromolecular Structural Analysis p. 433.
[50] M. Lingenfelder. Chiral recognition and supramolecular self-assembly of adsorbed amino acids and dipeptides at the submolecular level. *PhD thesis*, École Polytechnique Fédèrale de Lausanne; **2008**.
[51] D. A. Egger, Z. F. Liu, J. B. Neaton, L. Kronik, *Nano Lett.* **2015**, *15*, 2448.
[52] Z. F. Liu, D. A. Egger, S. Refaely-Abramson, L. Kronik, J. B. Neaton, *J. Chem. Phys.* **2017**, *146*(9), 092326.
[53] S. M. Barlow, K. J. Kitching, S. Haq, N. V. Richardson, *Surface Sci.* **1998**, *401*(3), 322.
[54] S. M. Barlow, S. Louafi, D. Le Roux, et al., *Langmuir* **2004**, *20*(17), 7171.
[55] C. Méthivier, V. Humblot, C. M. Pradier, *Surface Sci.* **2015**, *632*, 88.
[56] E. M. Marti, S. M. Barlow, S. Haq, R. Raval, *Surface Sci.* **2002**, *501*(3), 191.
[57] E. M. Marti, A. Quash, C. Methivier, P. Dubot, C. M. Pradier, *Colloids Surf. A: Physicochem. Eng. Aspects* **2004**, *249*(1), 85.
[58] T. Eralp, A. Shavorskiy, Z. V. Zheleva, G. Held, N. Kalashnyk, Y. Ning, T. R. Linderoth, *Langmuir* **2010**, *26*(24), 18841.
[59] F. Iori, R. Di Felice, E. Molinari, S. Corni, *J. Comput. Chem.* **2009**, *30*(9), 1465.
[60] L. B. Wright, P. M. Rodger, S. Corni, T. R. Walsh, *J. Chem. Theory. Comput.* **2013**, *9*(3), 1616.
[61] H. Heinz, T. J. Lin, R. Kishore Mishra, F. S. Emami, *Langmuir* **2013**, *29*(6), 1754.
[62] S. Koslowski Scanning tunneling microscopy on large bio-molecular systems on surfaces. *PhD thesis*, École Polytechnique Fédèrale de Lausanne; **2017**.
[63] M. Todorović, M. U. Gutmann, J. Corander, P. Rinke, *Npj Comput. Mater.* **2019**, *5*(1), 35.
[64] S. E. Schönborn, S. Goedecker, S. Roy, A. R. Oganov, *J. Chem. Phys.* **2009**, *130*(14), 144108.
[65] J. M. Dieterich, B. Hartke, *Mol. Phys.* **2010**, *108*(3–4), 279.
[66] A. Supady, V. Blum, C. Baldauf, *J. Chem. Informat. Model.* **2015**, *55*(11), 2338.

[67] R. Tom, T. Rose, I. Bier, H. O?Brien, Á. Vázquez-Mayagoitia, N. Marom, *Comput. Phys. Commun.* **2020**, *250*, 107170.
[68] L. Hörmann, A. Jeindl, A. T. Egger, M. Scherbela, O. T. Hofmann, *Comput. Phys. Commun.* **2019**, *244*, 143.
[69] S. Olsson, H. Wu, F. Paul, C. Clementi, F. Noé, *Proc. Nat. Acad. Sci.* **2017**, *114*(31), 8265.
[70] G. Fraux, M. Ceriotti, *Chemiscope: an interactive structure/property explorer for materials and molecules*; https://chemiscope.org (accessed: April 2020).

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Maksimov D, Baldauf C, Rossi M. The conformational space of a flexible amino acid at metallic surfaces. *Int J Quantum Chem*. 2021;121:e26369. https://doi.org/10.1002/qua.26369