molecules

MDPI

Article

Microsolvation of a Proton by Ar Atoms: Structures and Energetics of Ar_nH^+ Clusters

María Judit Montes de Oca-Estévez [1,2,†] and Rita Prosmiti [1,*]

[1] Institute of Fundamental Physics (IFF-CSIC), CSIC, Serrano 123, 28006 Madrid, Spain; juditmontesdeoca@iff.csic.es
[2] Atelgraphics S.L., Mota de Cuervo 42, 28043 Madrid, Spain
* Correspondence: rita@iff.csic.es; Tel.: +34-91-5616800 (ext. 442292)
† Doctoral Programme in Theoretical Chemistry and Computational Modelling, Doctoral School, Universidad Autónoma de Madrid, 28049 Madrid, Spain.

Abstract: We present a computational investigation on the structural arrangements and energetic stabilities of small-size protonated argon clusters, Ar_nH^+. Using high-level ab initio electronic structure computations, we determined that the linear symmetric triatomic ArH^+Ar ion serves as the molecular core for all larger clusters studied. Through harmonic normal-mode analysis for clusters containing up to seven argon atoms, we observed that the proton-shared vibration shifts to lower frequencies, consistent with measurements in gas-phase IRPD and solid Ar-matrix isolation experiments. We explored the sum-of-potentials approach by employing kernel-based machine-learning potential models trained on CCSD(T)-F12 data. These models included expansions of up to two-body, three-body, and four-body terms to represent the underlying interactions as the number of Ar atoms increases. Our results indicate that the four-body contributions are crucial for accurately describing the potential surfaces in clusters with $n > 3$. Using these potential models and an evolutionary programming method, we analyzed the structural stability of clusters with up to 24 Ar atoms. The most energetically favored Ar_nH^+ structures were identified for magic size clusters at $n = 7$, 13, and 19, corresponding to the formation of Ar-pentagon rings perpendicular to the ArH^+Ar core ion axis. The sequential formation of such regular shell structures is compared to ion yield data from high-resolution mass spectrometry measurements. Our results demonstrate the effectiveness of the developed sum-of-potentials model in describing trends in the nature of bonding during the single proton microsolvation by Ar atoms, encouraging further quantum nuclear studies.

Keywords: ab initio electronic structure calculations; molecular interactions; machine learning potentials; noble gas proton-bound clusters; microsolvation structures

Citation: Montes de Oca-Estévez, M.J.; Prosmiti, R. Microsolvation of a Proton by Ar Atoms: Structures and Energetics of Ar_nH^+ Clusters. *Molecules* **2024**, *29*, 4084. https://doi.org/10.3390/molecules29174084

Academic Editor: Chao Dong

Received: 27 July 2024
Revised: 22 August 2024
Accepted: 24 August 2024
Published: 28 August 2024

1. Introduction

Small-size noble gas hydride cations have aroused recent interest [1–8], as they could help enhance knowledge about the stability and behavior of the new systems discovered in the interstellar medium (ISM), such as the HeH^+ [9] and ArH^+ [10]. Consequently, they offer a strong starting point [11–18] for investigating the formation of higher-order complexes, when solvated by more than one noble gas (Ng) atom. Solvent species may interact strongly with the dissolved ions, forming a solvation shell around the charged center [19,20], with the number of solvent atoms needed for shell closure depending on the balance of the underlying interactions.

From a computational point of view, such $[Ng_nH]^+$ aggregates are extremely challenging systems, as quantum anharmonic treatments are required for studying their fully coupled dynamics and spectroscopy, and recent studies have been reported for the $n = 2$ and 3 cases [2,21–25]. They are known as proton-bound complexes (PBCs), which are characterized by being exceptionally bright molecules in the infrared (IR) region, with a

high intensity of the stretching vibrational mode of the shared central proton [26–28]. This fact makes them really easy targets to detect experimentally; thus, in recent years, there has been a revival of interest in studying these noble gas hydride cations by advanced computational quantum chemistry and experimental laboratory techniques, aiming to understand the chemical nature of these compounds and their evolution in the ISM [1].

From the experimental point of view, the detection of the Ar_2H^+ PBC was first reported in solid Ar/Kr matrix isolation spectroscopy experiments [29,30], and later on through electron bombardment matrix isolation of Ng/methanol mixtures [31], characterized by absorption bands at 905 and 903.8 cm^{-1}, respectively, while recently, Duncan and coworkers [8] have recorded the first gas phase infrared photodissociation (IRPD) spectra for the Ar_nH^+ (n = 3–7) clusters. These spectra show a series of strong bands, with those of the lowest frequency shifted gradually from 989 to 965 cm^{-1} as the size of the cluster increases from n = 3 to 7, respectively. Additionally, in this latter study [8], the geometry optimizations at Møller–Plesset perturbation, MP2/AVTZ, level of theory, for all clusters up to n = 7. They have revealed that the [ArH^+Ar] core structure remained practically invariant as more Ar atoms were added to the complex, in line with recent results from a coupled cluster with single, double, and perturbative triple excitations, CCSD(T)/AVTZ, and calculations on the structure and vibrational spectra of Ar_3H^+ [2]. More recently, high-resolution mass spectroscopy measurements have been reported on protonated and pristine He and Ar clusters of up to 50 noble gas atoms [3–5], indicating significant differences between the heaviest Ar and He protonated species, identifying magic sizes for the Ar_nH^+ for n = 7, 13, and 19.

In this context, the present work aims to characterize small finite-size Ar_nH^+ clusters, up to 24 Ar atoms. As such clusters consist of Ar atoms weakly bound to the covalent centro-symmetric Ar-H^+-Ar core, we have first investigated the interactions of the linear [ArH^+Ar] trimer with an additional Ar atom, performing high-level ab initio coupled cluster explicit correlated F12, CCSD(T)-F12/AVQZ, calculations. Both the configuration and energy datasets of these calculations have proven fundamental in training a machine learning (ML) potential energy surface (PES) model by employing a 2D reproducing kernel Hilbert space (RKHS) scheme [32,33]. In turn, we explored various simplified sum-of-potentials approaches, such as those including expansion up to two-body (2B), three-body (3B), and four-body (4B) terms, to represent the Ar-doped proton groups as the number of Ar atoms increases. Such analytical potential models were validated against ab initio MP2, CCSD(T), and CCSD(T)-F12 data using AVQZ and AV6Z basis sets [34–37]. The structural characteristics and energetics of the Ar_nH^+ clusters were investigated by potential optimizations through an evolutionary programming (EP) algorithm [38,39]. Comparisons with available data in the literature were discussed, contributing to confirm certain effects during the proton microsolvation by Ar atoms.

2. Computational Details, Results and Discussion

2.1. Binding and Structuring in Ar_nH^+ Clusters from Electronic Structure Calculations

As a starting point in exploring how the Ar atoms are arranged around the proton in the higher-order Ar_nH^+ clusters, geometry optimizations and frequency analysis were performed at CCSD(T)/AVQ/6Z, CCSD(T)-F12/AVQZ, and MP2/AVQZ levels of theory, depending on the size of the cluster. All ab initio electronic structure calculations were performed with MOLPRO 2022 program [40], while the DENEB software package [41] was employed to generate and organize all input and output data files, respectively.

Various initial guess configurations, some of them given in previous studies on such systems [24,42], were considered in order to identify the most stable conformer for each cluster. Given that the number of electrons involved increases substantially when a new Ar atom is added, we have considered performing the optimization calculations at the MP2 level of theory with the AVQZ basis set.

Figure 1 and Table S1 (see the Supplementary Materials) present our results on optimized structures derived from the MP2/AVQZ and CCSD(T)/AV6Z or CCSD(T)-F12/AVQZ calculations performed on the Ar_nH^+ clusters, and the binding energies with respect the dissociation of one Ar atom from each cluster, $Ar_nH^+ \rightarrow Ar_{n-1}H^+ + Ar$. By examining the corresponding configurations, one can observe a common core structure in each Ar_nH^+ cluster, consisting of a central linear [ArH^+Ar], that remains almost unchanged as additional Ar atoms are added to the cluster. One can note that the first two Ar atoms form a relatively strong bond of about 5500 cm^{-1}, compared to the rest of the argons that bond more weakly, with energies of 600–900 cm^{-1} (see Table S1). Similarly, the bond distances of the core are notably shorter than in the rest of the unions (1.5 Å vs. 3.2/4.4 Å), as seen in Figure 1. Ar atoms outside the core align perpendicularly to the molecular axis of [ArH^+Ar] until the complex reaches a size of n = 7, where the first solvation shell is completed, forming a pentagonal dipyramid structure. For larger clusters with n = 8 and 9, the Ar atoms are accommodated at significantly larger distances, around 4.4 Å from the core. From such initial analysis of the structural characteristics of these cations, it is clear that the binding in the proton-bound dimer is strong, while additional Ar atoms form weakly bound clusters. This indicates that the interactions between this cationic core and a single Ar atom could represent the most dominant contributions in the description of the PESs of higher-order Ar_nH^+ clusters. With this in mind, the next step entails examining the intermolecular interactions in the Ar_3H^+.

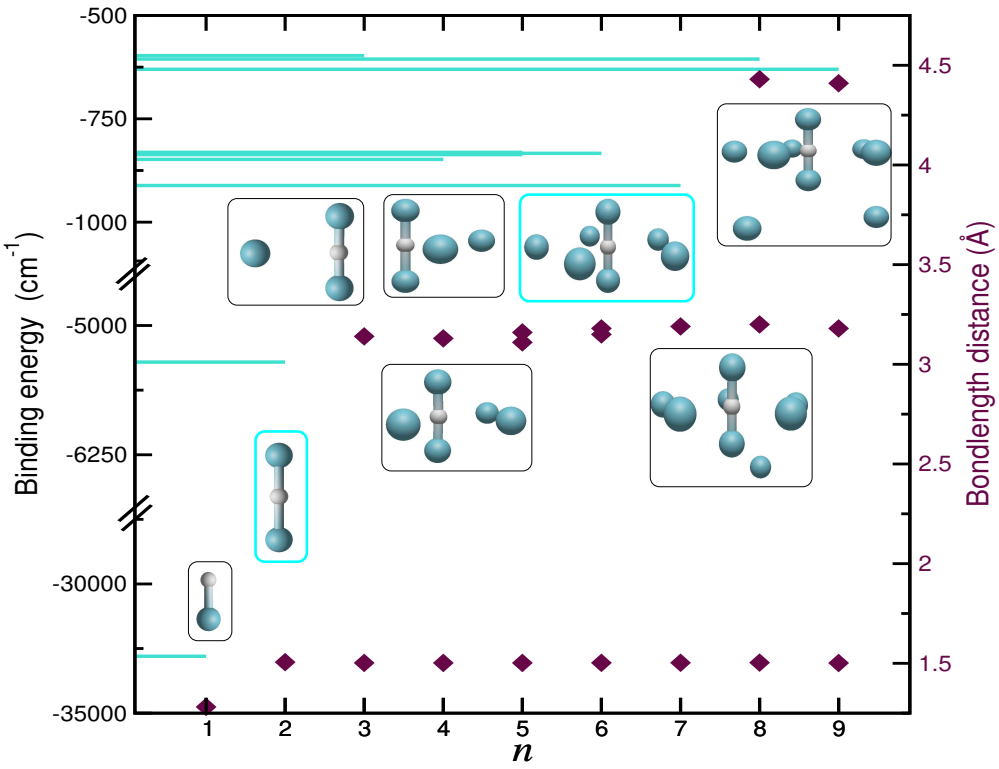

Figure 1. Binding energies and bondlengths of the optimized structures (see inset plots) of the Ar_nH^+ complexes (with n = 1–9) from CCSD(T)/AV6Z (n = 1–2) and MP2/AVQZ (n = 3–9) calculations.

2.2. Building Up a ML PES for the Ar_3H^+

The generation of high-quality data that represent the entire configurational space is a fundamental process in the construction of any PES. Here, we used Jacobi coordinates (r, R, θ) to define the Ar_3H^+ system (see inset plot in the upper panel of Figure 2), where r is the vector along the Ar–H^+–Ar distance, R is the vector along the distance between the Ar atom and the center of mass of the [ArH^+Ar] core, and θ is the angle between the (**r**, **R**) vectors. Since the structure of the [ArH^+Ar] core remains nearly unchanged as the size of the Ar_nH^+ clusters increases, we have decided to keep the r coordinate fixed to the

equilibrium bond length of the ArH^+Ar at r_e=3.0112 Å. In this way, the configuration space is sampled using the R and θ descriptors in the interval of R= 2.0–8.0 Å and θ= 0–90° (see Figure 2).

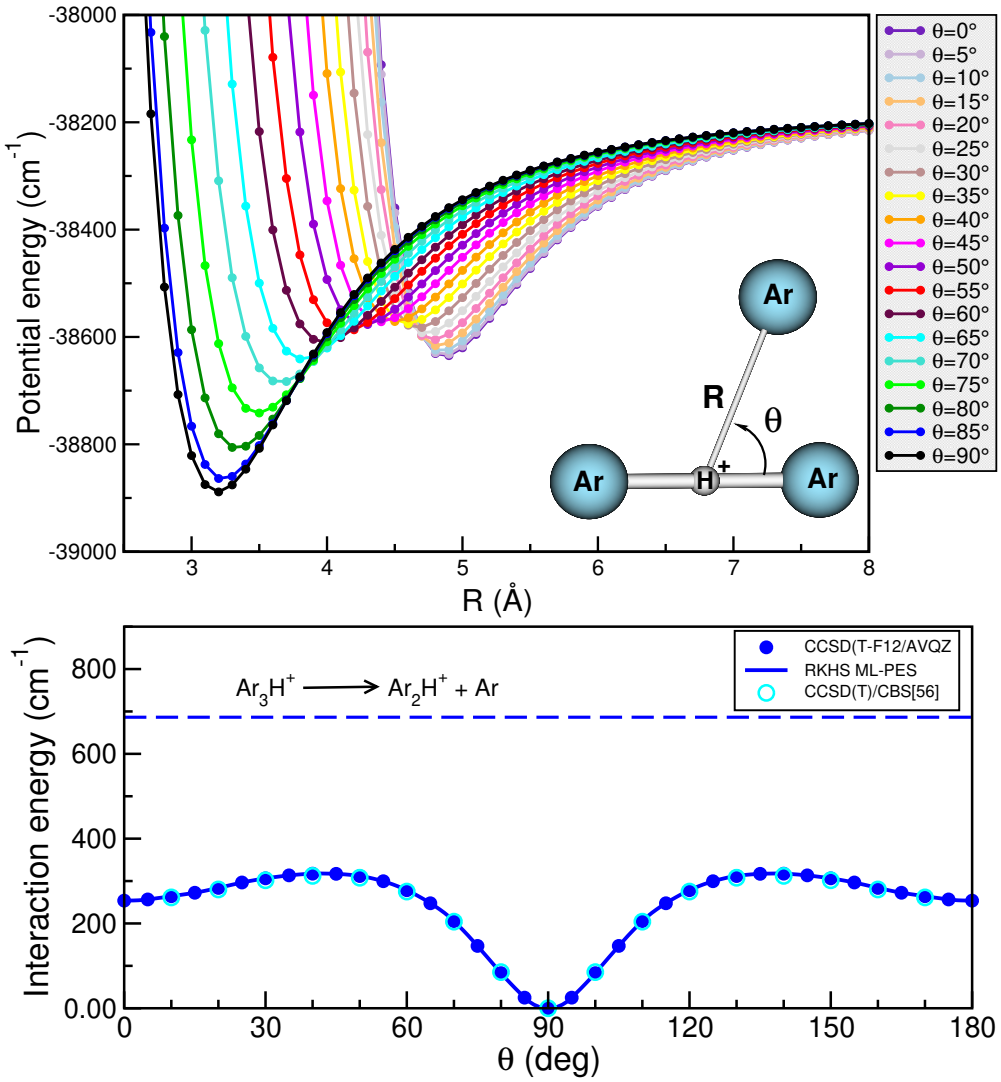

Figure 2. Interaction CCSD(T)-F12/AVQZ energies (see circle symbols) as a function of R distance at the indicated θ values (see **upper** panel) and minimum energy path along both R and θ coordinates as a function of θ angles (see **lower** panel) for the Ar_3H^+ cation (see inset plot). The corresponding RKHS ML-PES curves are also shown as solid lines, while the dissociation energy is plotted by long-dashed line.

Next, the performance of the electronic structure method employed is also an important issue in the PES construction. Thus, regarding the choice of the level of theory, and taking into account the increased demand of the computational resources, together with the outcome of previous benchmark studies of such complexes [24], a way to maintain the quality of the data is through the use of the CCSD(T)-F12 methodology employing the AVQZ basis set. As has been demonstrated in ref. [24], the CCSD(T)-F12 energies closely reproduce CCSD(T)/CBS[56] data. On this basis, the CCSD(T)-F12/AVQZ method was used for the present calculations, with the interaction energy defined as $\Delta E = E_{Ar_3H^+} - E_{Ar_2H^+} - E_{Ar} - \delta_{BSSE}$, with $E_{Ar_3H^+}$, $E_{Ar_2H^+}$, and E_{Ar} being the total energies of the Ar_3H^+ complex and Ar_2H^+ and Ar fragments, respectively, while δ_{BSSE} includes the basis set superposition error (BSSE) correction [43].

Among the ML-approaches, we have chosen to employ the RKHS method, which is a reasonable option due to its successful results with low dimensional systems and small training sets. Briefly, if $V(x)$ is the potential energy function with V_i corresponding to known potential energies at different molecular configurations x_i, then using the representer theorem for a general functional relationship, the $V(x)$ can be optimally approximated as a linear combination of suitable functions, $V(x) \approx V_i = \sum_{i=1}^{\mathcal{N}} C_i K(x, x_i)$, where C_i are coefficients and $K(x, x')$ is a kernel function. The coefficient C_i can be determined for all input $\mathbf{y} = [V_1, V_2, \cdots, V_{\mathcal{N}}]$ and x_i in the data set by solving the linear equation $\mathbf{C} = \mathbf{K}^{-1}\mathbf{y}$, with $\mathbf{C} = [C_1, \cdots, C_{\mathcal{N}}]$ being the vector coefficients, and $\mathbf{K}$ being the $\mathcal{N} \times \mathcal{N}$ reproducing kernel matrix. The explicit expression of multidimensional kernel functions $K(x, x')$ can be written as a direct product, $K(x, x') = \prod_{d=1-D} k(x, x')$, with D being the dimension of the multidimensional K kernel, and k are the one-dimensional (1D) kernels. Several 1D kernel functions for different purposes are available in the literature [44], and in this study, we have employed the $k_1^{n,m}$ and k_2 1D reproducing kernel functions for the distance-like (R) and angle-like (θ) variables [32], respectively, which combine kernels with physically motivated descriptors, like those determining the smoothness of the kernel function and its correct asymptotic behavior. The reduced coordinates are $x = R$ and $y = \cos\theta$; N_R and N_θ are the number of ab initio calculated points in each R and θ coordinate, with

$$k_1^{n,m}(x, x') = n^2 x_>^{-(m+1)} B(m+1, n)\, {}_2F_1(-n+1, m+1; n+m+1; \frac{x_<}{x_>}) \quad (1)$$

where $x_>$ and $x_<$ are the larger and smaller of the x and x', respectively, and

$$k_2(y, y') = \sum_{ll} \frac{(2ll+1)}{2} P_{ll}(y) P_{ll}(y') \quad (2)$$

The n and m superscripts refer to the order of smoothness of the function and its asymptotic behavior at large distances, with $n = 2$, and m = 3, as the R^{-4} accounts for the leading dispersion interaction between the Ar atom and ArH^+Ar molecular cation. B is the beta function, ${}_2F_1$ is the Gauss hyper-geometric function [32], and P_{ll} are the Legendre polynomials, with ll = 2k and k = 0–6. In turn, the potential form is given by $V(R, \theta; r_e) = \sum_{i=1}^{N_R} \sum_{j=1}^{N_\theta}, C_{ij} k_1^{n,m}(x_i, x) k_2(y_j, y)$, where the C_{ij} coefficients are obtained by solving the linear equations, with V(R_i, θ_k;r) being the ab initio CCSD(T)-F12/AVQZ energy at each (R_i, θ_j, r_e) grid point.

We have generated a total of 2679 ab initio CCSD(T)-F12/AVQZ datasets, with 1159 of them used in the training step, and 1520 in the testing step. The training points were equally distributed across a 61-point grid in the internuclear distance R, ranging from 2.0 to 8.0 Å, and a 19-point grid was used in the angle θ coordinate between 0 and 90°. We have employed a hold-out cross-validation scheme to select the best performing RKHS PES among six models trained on 200, 380, 400, 760, 610, and 1159 points, respectively. These points were selected with variable number of 20, 40, or 61 grid points in the R coordinate, and 10 or 19 in the θ coordinate, covering the whole radial and angular intervals. Figure 3 depicts the histograms of these training data sets and, as one can see, the energy distributions look quite similar, although by increasing the number of angular grid points permits a broader sampling, covering the whole energy range.

Next, in the left panel of Figure 4, we display histograms of the energy distribution of the 1520 CCSD(T)-F12/AVQZ configurations used for the PES evaluation during the testing process. In the right panel of Figure 4, we show the root-mean-squared error (RMSE) values as a function of the training dataset size. One can observe the systematic improvement of the RMSE of the RKHS ML-PES models as the size of the training dataset increases. The RMSE values were computed by averaging over the total number of 1520 testing data and over 506 randomly chosen points, as well as over 1341 and 445 randomly chosen test data with energies below the dissociation threshold. The quality of the six RKHS ML-PES models as a function of the size of the training data with 200, 380, 400, 760, 610,

and 1159 points is demonstrated in the right panel of Figure 4. As the size of the dataset increases, the corresponding RMSE value decreases. We also observed a large sensitivity in the RMSE values with respect to the sampling in the θ coordinate. In particular, we found that the RKHS ML-PES models trained on the 200 (20 $\times$ 10), 400 (40 $\times$ 10), and 610 (61 $\times$ 10) datasets show much higher RMSE values compared to those obtained from the 380 (20 $\times$ 19), 760 (40 $\times$ 19), and 1159 (61 $\times$ 19) ones. Additionally, we found that the calculated RMSE values for all six RKHS ML-PES trained models with respect to the total 1520 data, 1341 datasets with energies below the dissociation threshold, and the two corresponding 506/445-randomly chosen testing sets exhibit the same behavior.

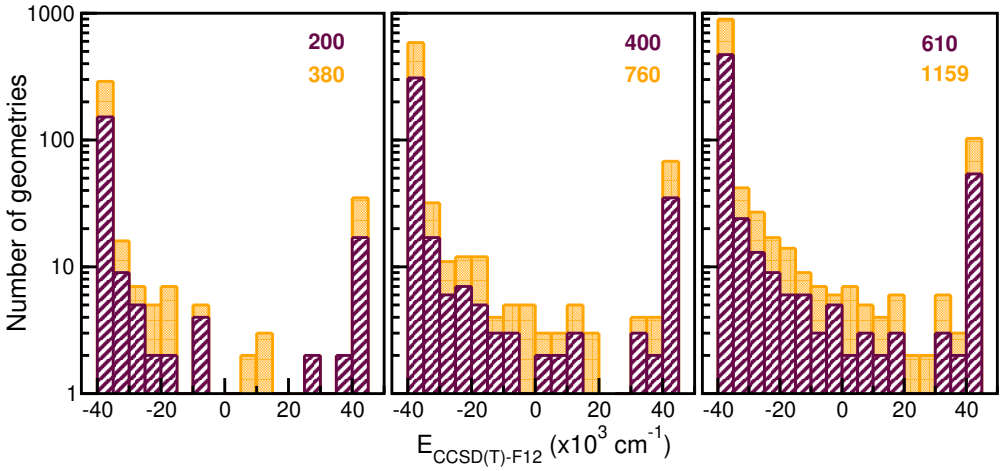

Figure 3. Histograms of training datasets employed in the hold-out cross-validation scheme for assessing the RKHS ML-PES models. The number of total configurations considered is given in the top of each panel. The left, middle and right side panels correspond to increasing numbers of grid points in R, 20, 40 and 61, respectively, while the orange and maroon colors show datasets with 10 and 19 grid points in θ.

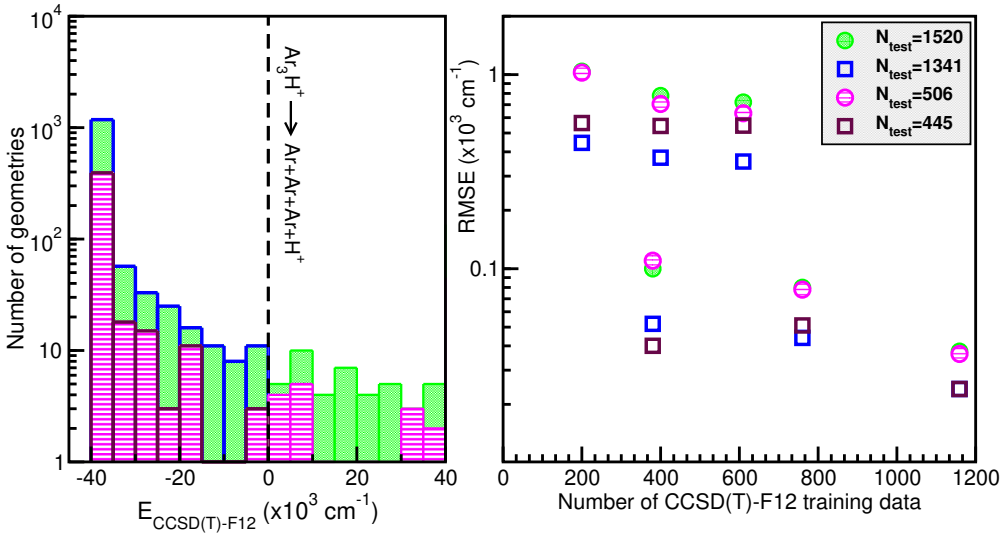

Figure 4. Histograms of the total and randomly generated testing datasets (see text) employed in hold-out cross-validation scheme for the RKHS ML-PES (**left** panel). The corresponding RMSE values of the RKHS ML-PES models as a function of the training set size (**right** panel). The black dashed line indicates the dissociation threshold.

Correlation plots are shown in Figure 5, demonstrating the performance of the chosen RKHS ML-PES model, which corresponds to that trained to 1159 CCSD(T)-F12, data against the energies of both training (see left panel) and testing (see right panel) data in both the

attractive and repulsive regions of the potential. The RMSE values are also displayed as a function of energy ranges (see inset plot in the right panel of Figure 5). One can see that the RMSE values outside the training zone are 24 cm^{-1} up to dissociation energies over 1341 configurations (see also the left panel of Figure 4), and a MAE lower than 0.07% over 179 configurations and up to energies of 600 cm^{-1} above dissociation.

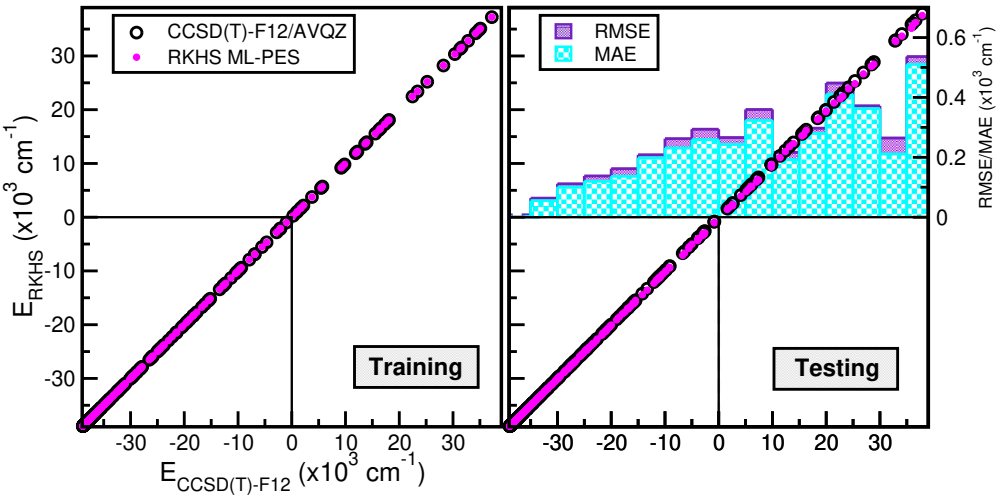

Figure 5. Correlation plots of the RKHS ML-PES model against the reference CCSD(T)-F12/AVQZ energies for both training (**left** panel) and testing (**right** panel) data. The corresponding average RMSE and MAE values along energy are also plotted.

Once we have constructed the 2D RKHS ML-potential for the Ar_3H^+ cation, we then analyzed the specific characteristics of its topology. To facilitate this analysis and ensure the smoothness of the PES, we present 1D curves (see Figure 2) and a 2D contour plot (see Figure 6).

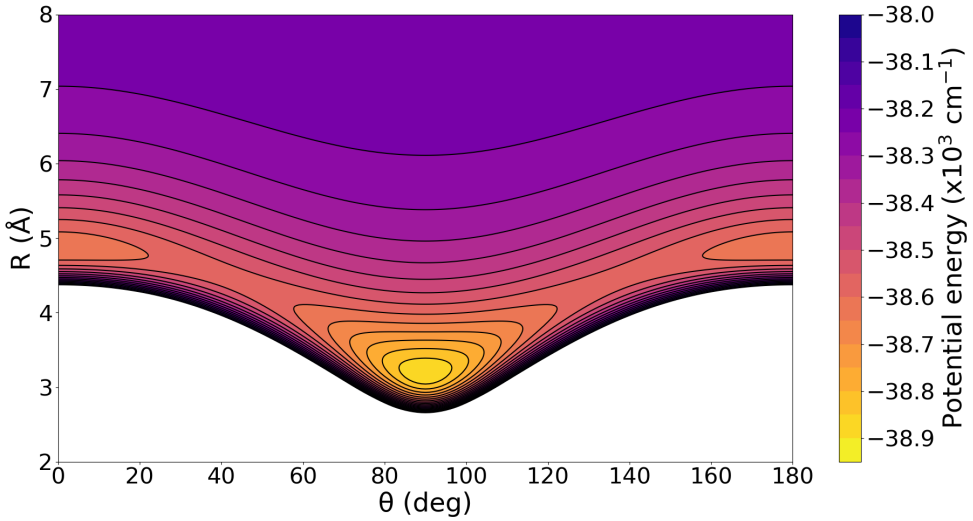

Figure 6. 2D contour plot of the RKHS ML-PES for Ar_3H^+ complex in the (θ,R)-plane. The equipotential curves are at energies of −39,000 to −38,000 cm^{-1} in intervals of 100 cm^{-1}.

One can clearly observe the presence of two minima on the Ar_3H^+ PES. The global minimum corresponds to the T-shaped configuration with a well-depth of 38,888.67 cm^{-1} for the RKHS ML-PES and CCSD(T)-F12/AVQZ data, located at R = 3.20 Å (0.06 Å larger than that predicted by the MP2/AVQZ optimization calculation; see Figure 1). The negligible difference between the ab initio data and the present RKHS ML-PES suggests that the RKHS model describes accurately and smoothly this region of the potential surface. The

second minimum lies around 250 cm^{-1} above the global one, and corresponds to the linear geometry of the complex, Ar-Ar-Ar-H^+, at an energy of −38,634.80 cm^{-1} with R = 4.90 Å. Again, there is an excellent agreement of these PES values compared to those from the ab initio CCSD(T)-F12 data at −38,655 cm^{-1} and 4.93 Å, respectively.

2.3. Modeling the Interaction in Higher-Order Ar_nH^+ Clusters

In the ongoing effort to characterize and understand the nature of interactions within higher-order size Ar_nH^+ clusters, we have employed the many-body expansion formalism [45], which is a fundamental approach for calculating the total energy (E_n) of systems consisting of n-bodies or monomers as

$$E_n = \sum_n E_i^{(1B)} + \sum_n E_{ij}^{(2B)} + \sum_n E_{ijk}^{(3B)} + \ldots + E^{(nB)} \quad (3)$$

where each term in the sum series reflects the contribution of specific interactions between n-bodies, such as the pairwise two-body interaction, $E_{ij}^{(2B)}$, and the higher nonadditive three-body, $E_{ijk}^{(3B)}$, up to n-body, $E^{(nB)}$, terms. Through such decomposition, investigation on complex systems are significantly simplified, and could provide a solid foundation for the analysis of physical and chemical phenomena. In this context, we will explore the capability of three analytical schemes based on the sum of two (2B)-, three (3B)-, and four-body (4B) interactions to represent and model the underlying interactions in Ar_nH^+ clusters. The corresponding expressions are given by

$$V_{2B}(\vec{R}_i) = \sum_{i=1-N} V_{[ArH]_i^+}(\vec{R}_i) + \sum_{i,l=1-N,l>i} V_{ArAr}(\vec{R}_{il}) \quad (4)$$

$$V_{3B}(\vec{R}_i, r) = \sum_{i=1-N} V_{[Ar_2H]_i^+}(\vec{R}_i, r) + \sum_{i,l=1-N,l>i} V_{ArAr}(\vec{R}_{il}) \quad (5)$$

$$V_{4B}(\vec{R}_i, r_e) = \sum_{i=1-N} V_{[Ar_3H]_i^+}(\vec{R}_i, r_e) + \sum_{i,l=1-N,l>i} V_{ArAr}(\vec{R}_{il}) \quad (6)$$

where the vectors R_i connect the proton with the i-th Ar atom, with polar and azimuthal angles θ_i and ϕ_i, respectively, while $\mathbf{r}$ and $\mathbf{r}_e$ are the vectors along the ArH^+ and $AHAr^+$ in Equations (5) and (6), respectively (see left and right panels of Figure 7, respectively), and R_{il} vectors connect the i-th and l-th Ar atoms in each Ar_nH^+ cluster (see Figure 7). The corresponding $V_{ArH^+}(\mathbf{R})$ and $V_{Ar_2H^+}(\mathbf{R}, \mathbf{r})$ terms are the RKHS ML-PESs of ArH^+ and Ar_2H^+ from refs. [24] and [46], respectively, while the $V_{Ar_3H^+}(\mathbf{R}, \mathbf{r_e})$ corresponds to the present Ar_3H^+ RKHS ML-PESs. The V_{ArAr} is the potential for Ar_2 from ref. [47].

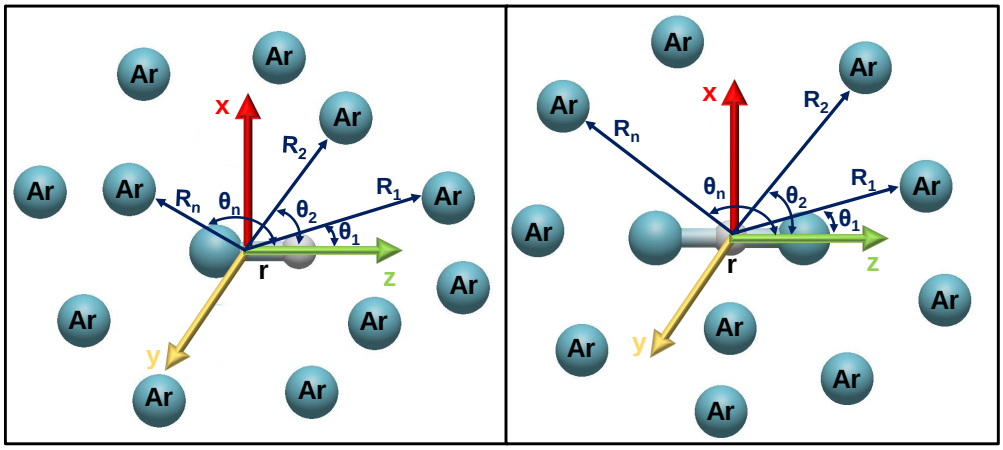

Figure 7. Coordinate system used for the Ar_nH^+ clusters in the sum-of-three-body (**left** panel) and sum-of-four-body (**right** panel) potential approaches.

2.4. Assessing the Sum-of-Potentials Models

In order to validate the three different sum-of-potential approaches adapted here (V_{2B}, V_{3B}, and V_{4B}), we have also carried out additional CCSD(T)-F12/AVQZ calculations for the pentatomic Ar_4H^+ cluster.

Figure 8 displays the potential energy curves obtained from the different sum-of-potentials analytical expressions (see Equations (4)–(6)), along with the corresponding ab initio computed interaction energies at CCSD(T)/CBS[56] for the n = 2 cluster along the R_1 coordinate, and CCSD(T)-F12/AVQZ for the n = 3 and 4 clusters along the R_2 and R_3/R_4 coordinates, respectively. As illustrated in the upper panels of Figure 8, we explored the cases with the ArH^+ r value fixed at its equlibrium 1.5056 Å distance, and thus the ArH^+Ar core r value is 3.0112 Å. In the n = 4 case, we considered selected configurations with fixed r and R_3 bondlengths at 3.0112 Å and 3.12 Å, respectively, or R_3 = R_4 with Ar atoms in T-shaped geometries (as shown in the inset plots in the lower panels).

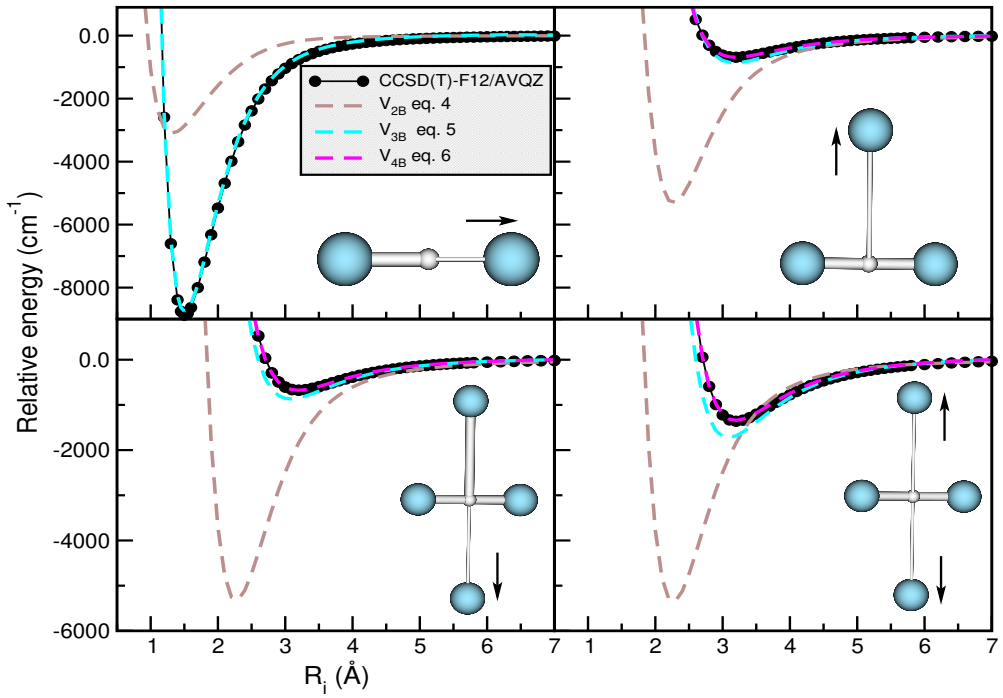

Figure 8. Potential curves obtained for the Ar_nH^+ (with n = 2, 3 and 4) clusters using the sum-of-potential approaches of Equations (4)–(6) (see color dashed lines), together with the calculated CCSD(T)/CBS[56] (for n = 2) and CCSD(T)-F12/AVQZ (for n = 3 and 4) interaction energies (black circles) as a function of indicated R_i coordinates.

One can see that for all the cases studied, the form of additive atom–atom interactions (V_{2B}) predicts potential energies that differ by around 5000 cm^{-1} in the well-depth region, as compared to the corresponding ab initio CCSD(T)/CBS[56] data. This fact reveals that three-body interactions will be important in the construction of Ar_2H^+ PESs, as confirmed by the results obtained when one employs the V_{3B} sum-of-potential form. In this case, the differences in that region are significantly reduced, counting 200 cm^{-1} for Ar_3H^+ and Ar_4H^+, as seen in the upper-right and lower-left panels of Figure 8, respectively. On the other hand, in the second configuration of Ar_4H^+ (see the lower-right panel in Figure 8), the difference increases up to 400 cm^{-1}. Regarding the V_{4B} sum-of-potential model, we observe that deviations between interaction energies are not greater than 20 cm^{-1} in both Ar_4H^+ configurations. For that reason, we choose the V_{4B} model, based on the sum of the CCSD(T)-F12/AVQZ $V_{Ar_3H^+}$ interactions, for all Ar_nH^+ clusters with $n > 3$ in the present study.

2.5. Searching for Ar_nH^+ Potential Energy Minima and Microsolvated Structures

Once the performance of the sum-of-potential approach was assessed, we proceed with searching for minimum energy structures on the V_{4B} potentials for each Ar_nH^+ cluster, and compare them with those obtained from the MP2/AVQZ calculations. Geometry optimization calculations on high-dimensional potential surfaces is a rather challenging task as the size of the molecular system increases. Thus, we employed an evolutionary programming (EP) algorithm [38] to localized the lowest energy structures for each Ar_nH^+ cluster on the corresponding V_{4B} PES. Such EP algorithms have demonstrated their accuracy and efficiency in solving numerical optimization problems in multi-dimensional space [38,39,48].

Briefly, the process begins by generating an initial population of $M = 100$ individuals for each Ar_nH^+ cluster under consideration, with n varying from 1 to 24. Each individual is characterized by a pair of real vectors (χ_i, η_i) for $i = 1$ to M, containing the Cartesian coordinates (χ_i) of all cluster atoms and their standard deviation (η_i) for Gaussian mutation (strategy parameter), governing the evolution of population dispersion over time. The initial coordinates χ_i, with $\eta_i = 1$, are randomly selected in the interval $(0, \Delta)$, where Δ is a displacement factor enhancing resolution. Each parent set (χ_i, η_i) evolves to generate a new population through mutation [38], while for each individual in the joint parent–child group $(2M$ individuals), q (tournament size equals 100) opponents are randomly chosen from the $2M - 1$ individuals for comparison. The individual with the lower potential energy in each encounter emerges victorious. The top M individuals then serve as parents for the next generation, and this process continues iteratively. Convergence is achieved when the potential energy difference between two consecutive generations falls below a threshold value of 10^{-3} cm^{-1}. Thus, by employing the EP optimization procedure for the potential surface of Equation (6), we were able to calculate the corresponding minima for the Ar_nH^+ clusters. The obtained results are summarized in Figure 9 and Table S1 in comparison with the MP2/AVQZ and CCSD(T)-F12 values (see also in Figure 1), as well as with previously data available in the literature [8]. Both set of data consistently demonstrate good agreement, with differences not exceeding 30 cm^{-1} in most cases, corresponding to errors of around 5%. In Figure 9, we also included zero-point energy (ZPE) corrections, $\mathcal{E}_{\mathrm{ZPE}}$, in the MP2/AVQZ energy values, by treating each cluster in the framework of the harmonic approximation. The corresponding ZPE values are obtained from normal-mode frequency calculations at each cluster's optimized structure, and are found to significantly affect the energetics of the clusters, although one should consider such values as an upper limit, taking into account the contribution of anharmonities.

In this context, Figure 10 presents comparisons for the fundamental asymmetric Ar–H^+ stretching (or proton-shared) vibration frequencies between the harmonic normal mode values at MP2/AVQZ level of theory and the vibrational bands observed in the experimentally recorded gas phase IRPD spectra [8] as the size of the Ar_nH^+ clusters increases from n = 2 to 7, as well as with that from solid Ar-matrix isolation measurements [29]. At first glance, one can clearly observe that the proton-shared harmonic frequency is shifted to lower values as the cluster size increases, in accord with the behavior observed in the gas phase IRPD spectra. However, the frequencies from the harmonic normal-mode analysis are found to be higher by 75–82 cm^{-1} than the IRPD reported values for n = 3 to 7 clusters [8], and also by 95 cm^{-1} compared to the anharmonic value reported by quantum calculations in the n = 2 case [24]. We should further notice that the gradual shift of the proton stretch vibration to lower frequencies as the cluster size increases reaches a limited value near to 1044 or 965 cm^{-1} from the harmonic approach and IRPD experiment, respectively. Such finding indicates the first shell closing at n = 7 in the gas phase clusters, although by comparing these frequency values with that observed in solid Ar matrix experiments at 905 cm^{-1} seems that additional shells and different local arrangements might also have an influence.

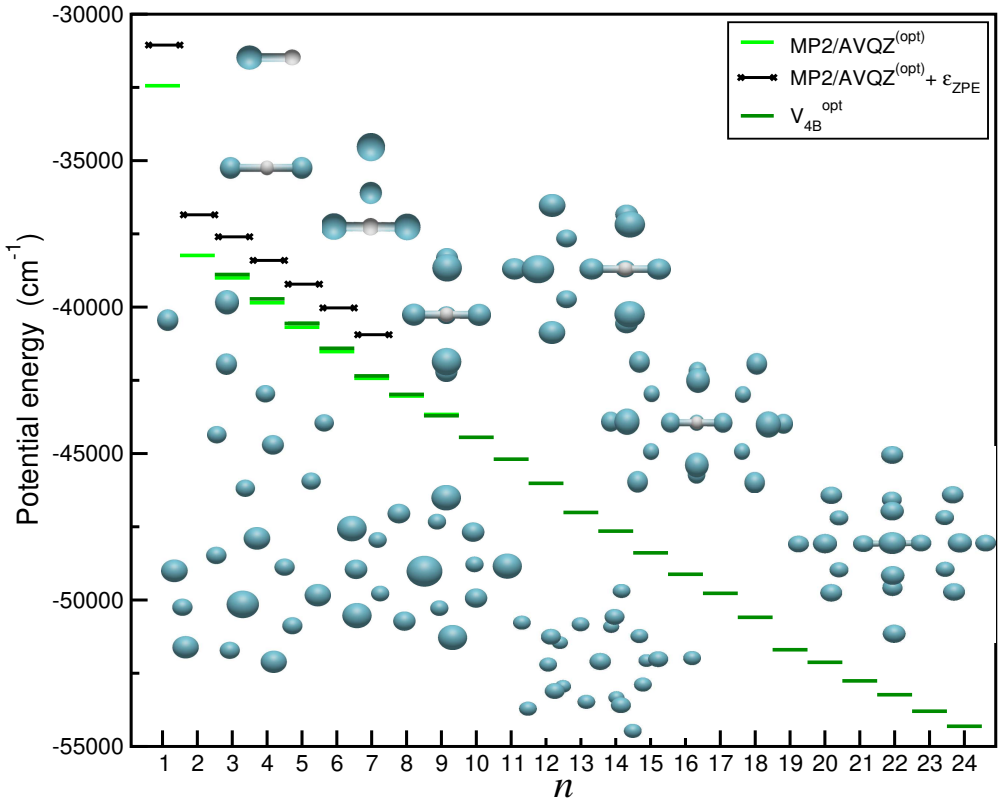

Figure 9. Schematic representation of selected optimal low-lying structures and their energetics obtained using the V_{4B} approach of Equation (6) and the GA algorithm for the Ar_nH^+ clusters. Energy values from the MP2/AVQZ optimizations and their ZPE corrections, $\mathcal{E}_{ZPE}$, given by the corresponding harmonic approximation, are also shown.

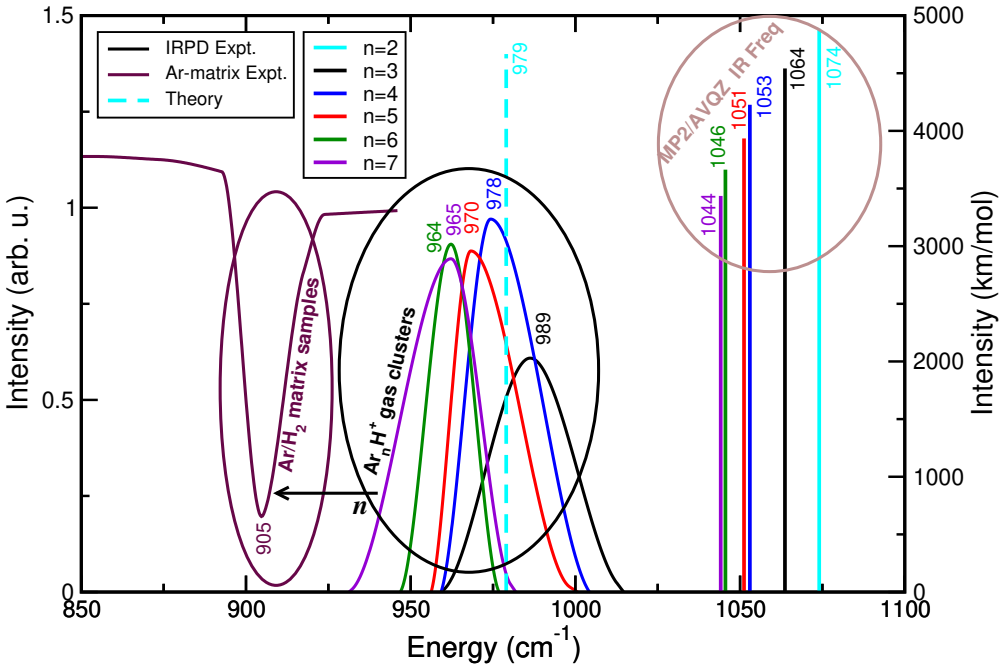

Figure 10. Computed harmonic frequencies for the indicated Ar_nH^+ clusters from MP2/AVQZ calculations in comparison with data from gas phase IRPD [8] and solid Ar-matrix isolation measurements [29]. The anharmonic fundamental frequency for Ar_2H^+ from ref. [24] is also displayed.

Finally, the stability of each cluster size was investigated concerning its nearest neighbors. In Figure 11, we illustrate the single-atom evaporative energies, defined as E_{n-1}–E_n,

and the average energy per added Ar atom, E_n/n (see inset plot). These quantities were determined through both EP optimizations using the V_{4B} potential, and MP2/AVQZ optimization calculations, and compared with previously reported dissociation energies for loosing a single Ar atom for the cationic cluster [5]. The dependence of the cluster size on the evaporation energy shows pronounced steps at n = 4 and 7, indicating the completion of the first solvation shell or the formation of a particularly stable structure. It is observed that results from both EP and ab initio calculations exhibit similar behavior starting from the n = 4 cluster. By examining the obtained E_n/n values, it is interesting to note that only the first two Ar atoms are independent and strongly bound to the H^+, yielding about −32,440 and −19,120 cm^{-1} E_n/n energies, respectively. For all sequential Ar atoms, this value is reduced, indicating a progressive importance of the Ar–Ar interaction. We found that the average energies per Ar atom approach a plateau as n increases, around −5000 cm^{-1} at n = 9. This behavior implies that the addition of new Ar atoms does not significantly alter the energetics of the larger clusters.

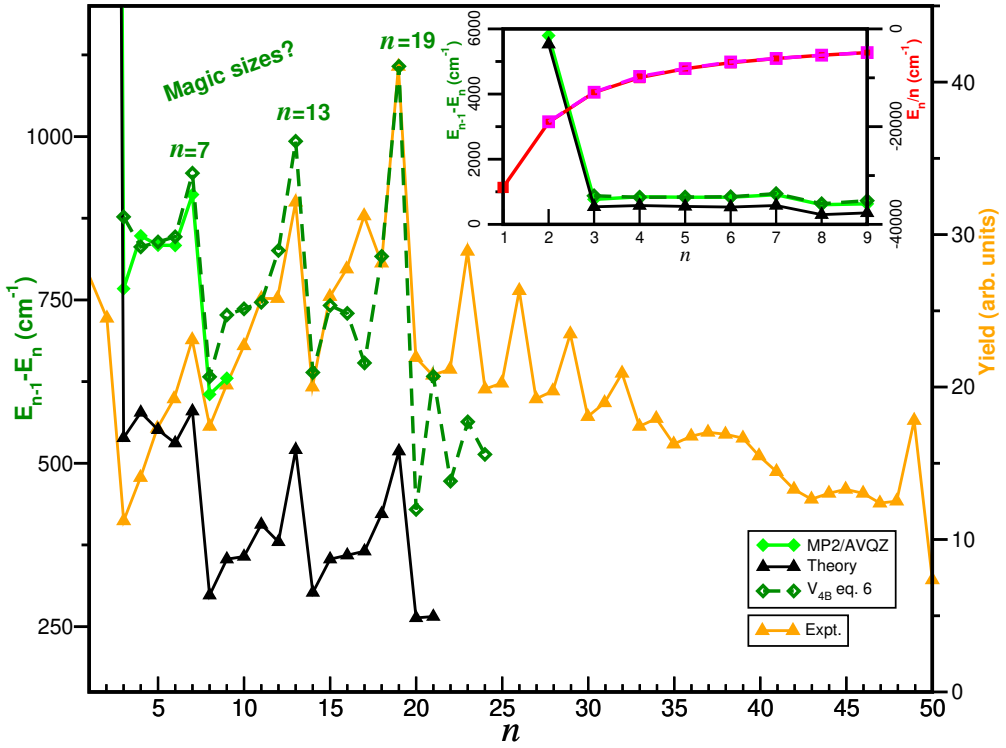

Figure 11. Computed single-atom evaporative energies (green color) of Ar_nH^+ clusters, their average energy per Ar atom (red color) from the V_{4B} potential EP optimizations (dashed lines) and MP2/AVQZ calculations (solid lines) as a function of n. Previously reported theoretical dissociation energies [5] and experimental ion yield values [3] are also displayed for comparison reasons.

Further, the computed evaporation energies are also compared (see in Figure 11 with experimental values of Ar_nH^+ ion yields up to 50 Ar atoms, as measured recently by high-resolution mass spectroscopy experiments [3]. By now examining the extracted distributions from ion yield measurements and the evaporation energies for the Ar_nH^+ clusters, one can see in Figure 11 the standout features corresponding to magic size clusters at n = 7, 13 and 19 in both data, as well as weaker peaks at n = 23, 26, 29 and 32 in the ion yield data. In Figure 9 and Table S1, we depict the computed optimal structures corresponding to these size clusters, and as one can clearly observe, the first solvation shell is built up by a central pentagon ring in the equatorial plane around the proton perpendicular to the ArH^+Ar core ion axis, while the 13-atom cluster features the formation of a second pentagon ring at the one end of the ArH^+Ar axis together with an additional Ar atom along it. The symmetrical accommodation of the remaining Ar atoms on the other

side of the central ring at distances of around 4.45 Å corresponds to the optimal $Ar_{19}H^+$ structure. On can also observe (see in Table S1) that for n = 20 starts a new shell, and as the cluster size increases the Ar atoms are accommodated accordingly.

3. Summary and Conclusions

We have investigated in detail the energetic and structural characteristics of Ar_nH^+ clusters (with n up to 24). As a starting point, optimizations were carried out on small-size cations using the MP2/AVQZ level of theory, and a common cationic [ArH^+Ar] core was observed in all of them. Such a finding has been verified by a normal-mode vibrational analysis showing red shifts in the asymmetric Ar-H^+ stretch frequency by increasing the cluster size from n = 2 to 7, in agreement with the gas phase IRPD experimental measurements. In order to address the interaction of the ArH^+Ar core with the addition of new Ar atoms, a two-dimensional (2D) ML-PES based on the RKHS method and trained on CCSD(T)-F12/AVQZ datasets has been constructed for the Ar_3H^+ complex. Next, analytical sum-of-potentials representations were generated through expansions up to two-body (V_{2B}), three-body (V_{3B}), and four-body (V_{4B}) interactions. These potentials were validated by direct comparisons with ab initio energies, noting that the V_{4B} form accurately predicts the CCSD(T)-F12 reference data for $n > 3$ clusters. Thus, such analytical potential expression has been then considered to describe the intermolecular interactions in the higher-order Ar_nH^+ clusters up to 24 Ar atoms.

The V_{4B} potential models employed in optimization EP calculations, and in comparison with the MP2/AVQZ results, have been provide information on the growth of the Ar_nH^+ clusters. According to the minimum energy structures obtained, the solvation of the proton occurs symmetrically around its two ends for small size clusters of up to 7 Ar atoms, forming a dipyramidal structure around the H^+. These results agree with the ab initio electronic structure calculations, suggesting that the present V_{4B} PESs allow extracting qualitative and quantitative knowledge about the energy and structural arrangements during the microsolvation processes of the H^+ cation in Ar atoms. In the future, it would be interesting to investigate the importance of quantum anharmonic, as well as temperature effects on the clusters structure and stability, and whether, for larger clusters, the optimized configurations remain symmetric or whether, on the contrary, the dominant influence of Ar-Ar interactions leads to more compact structures in subsequent solvation shells. Our studies could contribute to a comprehensive understanding of the energetic and structural characteristics of Ar_nH^+ clusters, as well as shedding light on interactions that take place in such systems during the proton's microsolvation by Ar atoms, and assign patterns observed experimentally.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/molecules29174084/s1, Table S1: Binding energies (in cm^{-1}) from CCSD(T)/AV6Z and MP2/AVQZ calculations for Ar_nH^+ complexes, with n = 1–3 and 1–9, respectively, $(V_{2B}/V_{3B})V_{4B}$ and previous theoretical work [8]. The corresponding optimized structures (side and top views) obtained from the CCSD(T)/AV6Z calculations for n = 1–3 and V_{4B} potential for n = 3–24 are also depicted.

Author Contributions: Conceptualization, R.P.; computations and software, M.J.M.d.O.-E.; formal analysis & validation, M.J.M.d.O.-E. and R.P.; writing—original draft preparation, M.J.M.d.O.-E. and R.P.; writing—review and editing, M.J.M.d.O.-E. and R.P.; visualization, M.J.M.d.O.-E. and R.P.; supervision, R.P.; funding acquisition, R.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by MCIN grant No. PID2020-114654GB-I00, Comunidad de Madrid grant No. IND2018/TIC-9467, CSIC-PEICT Ref: 2024AEP119, and COST Actions CA18212(MD-GAS), CA21101(COSY) and CA21126(NanoSpace).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data supporting reported results are available from the corresponding author upon request.

Acknowledgments: The authors acknowledge the Centro de Cálculo del IFF, SGAI (CSIC) and CESGA for allocation of computer time.

Conflicts of Interest: Author María Judit Montes de Oca-Estévez was employed by the company Atelgraphics S.L., Madrid, Spain. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Grandinetti, F. Cationic Noble-Gas Hydrides: From Ion Sources to Outer Space. *Front. Chem.* **2020**, *8*, 462. [CrossRef] [PubMed]
2. Tan, J.A.; Kuo, J.L. Structure and Vibrational Spectra of Ar_nH^+ (n = 2–3). *J. Phys. Chem. A* **2020**, *124*, 7726–7734. [CrossRef] [PubMed]
3. Lundberg, L.; Bartl, P.; Leidlmair, C.; Scheier, P.; Gatchell, M. Protonated and Cationic Helium Clusters. *Molecules* **2020**, *25*, 1066. [CrossRef]
4. Császár, A.; Szidarovszky, T.; Asvany, O.; Schlemmer, S. Fingerprints of microscopic superfluidity in HHe_n^+ clusters. *Mol. Phys.* **2019**, *117*, 1559–1583. [CrossRef]
5. Gatchell, M.; Martini, P.; Kranabetter, L.; Rasul, B.; Scheier, P. Magic sizes of cationic and protonated argon clusters. *Phys. Rev. A* **2018**, *98*, 022519. [CrossRef]
6. Stephan, C.J.; Fortenberry, R.C. The interstellar formation and spectra of the noble gas, proton-bound $HeHHe^+$, $HeHNe^+$ and $HeHAr^+$ complexes. *Mon. Not. R. Astron. Soc.* **2017**, *469*, 339–346. [CrossRef]
7. Fortenberry, R. Quantum astrochemical spectroscopy: Review. *Int. J. Quant. Chem.* **2017**, *117*, 81–91. [CrossRef]
8. McDonald, D.C.; Mauney, D.T.; Leicht, D.; Marks, J.H.; Tan, J.A.; Kuo, J.L.; Duncan, M.A. Communication: Trapping a proton in Argon: Spectroscopy and theory of the proton-bound Argon dimer and its solvation. *J. Chem. Phys.* **2016**, *145*, 231101. [CrossRef]
9. Güsten, R.; Wiesemeyer, H.; Neufeld, D.; Menten, K.; Graf, U.; Jacobs, K.; Klein, B.; Ricken, O.; Risacher, C.; Stutzki, J. Astrophysical detection of the helium hydride ion HeH^+. *Nature* **2019**, *568*, 357–359. [CrossRef]
10. Barlow, M.; Swinyard, B.; Owen, P.; Cernicharo, J.; Gomez, H.; Ivison, R.; Krause, O.; Lim, T.; Matsuura, M.; Miller, S.; et al. Detection of a Noble Gas Molecular Ion, $^{36}ArH^+$, in the Crab Nebula. *Science* **2013**, *342*, 1343–1345. [CrossRef]
11. Gianturco, F.; Filippone, F. Structure and dynamics of small protonated rare-gas clusters using quantum and classical methods. *Comput. Phys. Commun.* **2002**, *145*, 78–96. [CrossRef]
12. Giju, K.T.; Roszak, S.; Leszczynski, J. A theoretical study of protonated Argon clusters: Ar_nH^+ (n = 1–7). *J. Chem. Phys.* **2002**, *117*, 4803–4809. [CrossRef]
13. Ritschel, T.; Zülicke, L.; Kuntz, P.J. Cationic van-der-Waals Complexes: Theoretical Study of Ar_2H^+ Structure and Stability. *Z. Phys. Chem.* **2004**, *218*, 377–389. [CrossRef]
14. Ritschel, T.; Kuntz, P.J.; Zülicke, L. Structure and dynamics of cationic van-der-Waals clusters. *Eur. Phys. J. D* **2005**, *33*, 421–432. [CrossRef]
15. Montes de Oca-Estévez, M.J.; Prosmiti, R. Quantum computations in heavy noble-gas hydride cations: Reference energies and new spectroscopic data. *J. Mol. Graph. Model.* **2023**, *124*, 108562. [CrossRef]
16. Montes de Oca-Estévez, M.J.; Darna, B.; García-Ruíz, B.; Prosmiti, R.; González-Lezana, T.; Koner, D. Ar + ArH^+ Reactive Collisions of Astrophysical Interest: The Case of ^{36}Ar. *ChemPhysChem* **2023**, *24*, e202300450. [CrossRef] [PubMed]
17. Montes de Oca-Estévez, M.J.; Prosmiti, R. Automated learning data-driven potential models for spectroscopic characterization of astrophysical interest noble gas-containing NgH_2^+ molecules. *AIChem* **2024**, *2*, 100059. [CrossRef]
18. García-Vázquez, R.M.; Márquez-Mijares, M.; Rubayo-Soneira, J.; Denis-Alpizar, O. Relaxation of ArH^+ by collision with He: Isotopic effects. *A & A* **2019**, *631*, A86.
19. Pauling, L. *The Nature of the Chemical Bond and the Structure of Molecules and Crystals*; Cornell University Press: Ithaca, NY, USA, 1954.
20. Beyer, M.; Lammers, A.; Savchenko, E.V.; Niedner-Schatteburg, G.; Bondybey, V.E. Proton solvated by noble-gas atoms: Simplest case of a solvated ion. *Phys. Chem. Chem. Phys.* **1999**, *1*, 2213–2221. [CrossRef]
21. Fortenberry, R.C. Rovibrational Characterization and Interstellar Implications of the Proton-Bound, Noble Gas Complexes: $ArHAr^+$, $NeHNe^+$, and $ArHNe^+$. *ACS Earth Space Chem.* **2017**, *1*, 60–69. [CrossRef]
22. Tan, J.A.; Kuo, J.L. A theoretical study on the infrared signatures of proton-bound rare gas dimers (Rg–H^+–Rg), Rg = Ne, Ar, Kr, and Xe. *J. Chem. Phys.* **2019**, *150*, 124305. [CrossRef] [PubMed]
23. Koner, D. Quantum and quasiclassical dynamical simulations for the Ar_2H^+ on a new global analytical potential energy surface. *J. Chem. Phys.* **2021**, *154*, 054303. [CrossRef] [PubMed]
24. Montes de Oca-Estévez, M.J.; Valdés, A.; Prosmiti, R. A kernel-based machine learning potential and quantum vibrational state analysis of the cationic Ar hydride (Ar_2H^+). *Phys. Chem. Chem. Phys.* **2024**, *26*, 7060–7071. [CrossRef]
25. Montes de Oca-Estévez, M.J.; Valdés, A.; Koner, D.; González-Lezana, T.; Prosmiti, R. Quantum computations on a new neural network potential for the proton-bound noble-gas Ar_2H^+ complex: Isotopic effects. *Chem. Phys. Lett.* **2024**.

26. Botschwina, P.; Dutoi, T.; Mladenović, M.; Oswald, R.; Schmatz, S.; Stoll, H. Theoretical investigations of proton-bound cluster ions. *Faraday Discuss.* **2001**, *118*, 433–453. [CrossRef]
27. Fortenberry, R.C.; Yu, Q.; Mancini, J.S.; Bowman, J.M.; Lee, T.J.; Crawford, T.D.; Klemperer, W.F.; Francisco, J.S. Communication: Spectroscopic consequences of proton delocalization in $OCHCO^+$. *J. Chem. Phys.* **2015**, *143*, 071102. [CrossRef]
28. Yu, Q.; Bowman, J.M.; Fortenberry, R.C.; Mancini, J.S.; Lee, T.J.; Crawford, T.D.; Klemperer, W.; Francisco, J.S. Structure, Anharmonic Vibrational Frequencies, and Intensities of $NNHNN^+$. *J. Phys. Chem. A* **2015**, *119*, 11623–11631. [CrossRef] [PubMed]
29. Bondybey, V.E.; Pimentel, G.C. Infrared Absorptions of Interstitial Hydrogen Atoms in Solid Argon and Krypton. *J. Chem. Phys.* **1972**, *56*, 3832–3836. [CrossRef]
30. Milligan, D.E.; Jacox, M.E. Infrared spectroscopic evidence for the stabilization of HAr_n^+ in solid Argon at 14 K. *J. Mol. Spectrosc.* **1973**, *46*, 460–469. [CrossRef]
31. Fridgen, T.D.; Parnis, J.M. Electron bombardment matrix isolation of Rg/Rg/methanol mixtures (Rg = Ar, Kr, Xe): Fourier-transform infrared characterization of the proton-bound dimers Kr_2H^+, Xe_2H^+, $(ArHKr)^+$ and $(ArHXe)^+$ in Ar matrices and $(KrHXe)^+$ and Xe_2H^+ in Kr matrices. *J. Chem. Phys.* **1998**, *109*, 2155–2161. [CrossRef]
32. Ho, T.; Rabitz, H. A general method for constructing multidimensional molecular potential energy surfaces from ab initio calculations. *J. Chem. Phys.* **1996**, *104*, 2584–2597. [CrossRef]
33. Hollebeek, T.; Ho, T.S.; Rabitz, H. A fast algorithm for evaluating multidimensional potential energy surfaces. *J. Chem. Phys.* **1997**, *106*, 7223–7227. [CrossRef]
34. Dunning, T.H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023. [CrossRef]
35. Woon, D.E.; Dunning, T.H. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371. [CrossRef]
36. van Mourik, T.; Wilson, A.K.; Dunning, T.H. Benchmark calculations with correlated molecular wavefunctions. XIII. Potential energy curves for He2, Ne2 and Ar2 using correlation consistent basis sets through augmented sextuple zeta. *Mol. Phys.* **1999**, *96*, 529–547. [CrossRef]
37. Pritchard, B.P.; Altarawy, D.; Didier, B.; Gibsom, T.D.; Windus, T.L. A New Basis Set Exchange: An Open, Up-to-date Resource for the Molecular Sciences Community. *J. Chem. Inf. Model.* **2019**, *59*, 4814–4820. [CrossRef]
38. Iwamatsu, M. Applying evolutionary programming to structural optimization of atomic clusters. *Comput. Phys. Commun.* **2001**, *142*, 214–218. [CrossRef]
39. Fogel, D.B. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*; John Wiley and Sons: Hoboken, NJ, USA, 2006.
40. Werner, H.J.; Knowles, P.; Knizia, G.; Manby, F.; Schütz, M. MOLPRO, Version 2022.2. a Package of Ab Initio Programs. 2022. Available online: http://www.molpro.net (accessed on 15 February 2024).
41. DENEB 1.30 Beta: The Nanotechnology Software by Atelgraphics. Available online: http://www.atelgraphics.com (accessed on 30 January 2024).
42. Grabowski, S.J.; Ugalde, J.M.; Andrada, D.M.; Frenking, G. Comparison of Hydrogen and Gold Bonding in $[XHX]^-$, $[XAuX]^-$, and Isoelectronic $[NgHNg]^+$, $[NgAuNg]^+$ (X = Halogen, Ng = Noble Gas). *Eur. J. Chem.* **2016**, *22*, 11317–11328. [CrossRef] [PubMed]
43. Boys, S.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553–566. [CrossRef]
44. Unke, O.T.; Meuwly, M. Toolkit for the construction of reproducing kernel-based representations of data: Application to multidimensional potential energy surfaces. *J. Chem. Inf. Model.* **2017**, *57*, 1923–1931. [CrossRef]
45. Murrell, J.N. *Molecular Potential Energy Functions*; Wiley: Hoboken, NJ, USA, 1984.
46. Montes de Oca-Estévez, M.J.; Prosmiti, R. Computational Characterization of Astrophysical Species: The Case of Noble Gas Hydride Cations. *Front. Chem.* **2021**, *9*, 187. [CrossRef] [PubMed]
47. Tong, X.-F.; Yang, C.-L.; Xiao, J.; Wang, M.-S.; Ma, X.-G. Theoretical study on the complexes of He, Ne and Ar. *Chin. Phys. B* **2010**, *19*, 123102. [CrossRef]
48. Perez de Tudela, R.; Barragan, P.; Valdes, A.; Prosmiti, R. Energetics and solvation structure of a dihalogen dopant (I_2) in ^{4}He clusters. *J. Phys. Chem. A* **2014**, *118*, 6492–6500. [CrossRef] [PubMed]