King's Research Portal

Document Version
Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

Download date: 10. Jan. 2025

.

Mapping the conformational free energy of aspartic acid in the gas phase and in aqueous solution

Federico Comitani, Kevin Rossi, Michele Ceriotti, M. Eugenia Sanz, and Carla Molteni

---

Articles you may be interested in

Metadynamic metainference: Convergence towards force field independent structural ensembles of a disordered peptide
The Journal of Chemical Physics **146**, 165102 (2017); 10.1063/1.4981211

Reducing the number of mean-square deviation calculations with floating close structure in metadynamics
The Journal of Chemical Physics **146**, 115101 (2017); 10.1063/1.4978296

Defining and quantifying frustration in the energy landscape: Applications to atomic and molecular clusters, biomolecules, jammed and glassy systems
The Journal of Chemical Physics **146**, 124103 (2017); 10.1063/1.4977794

Cheap but accurate calculation of chemical reaction rate constants from ab initio data, via system-specific, black-box force fields
The Journal of Chemical Physics **147**, 161701 (2017); 10.1063/1.4979712

 Perspective: Dissipative particle dynamics
The Journal of Chemical Physics **146**, 150901 (2017); 10.1063/1.4979514

Learning molecular energies using localized graph kernels
The Journal of Chemical Physics **146**, 114107 (2017); 10.1063/1.4978623

Mapping the conformational free energy of aspartic acid in the gas phase and in aqueous solution

Federico Comitani,[1,a)] Kevin Rossi,[1] Michele Ceriotti,[2] M. Eugenia Sanz,[3] and Carla Molteni[1]

[1]*Department of Physics, King's College London, Strand, London WC2R 2LS, United Kingdom*
[2]*Laboratory of Computational Science and Modeling, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland*
[3]*Department of Chemistry, King's College London, Britannia House, 7 Trinity Street, London SE1 1DB, United Kingdom*

The conformational free energy landscape of aspartic acid, a proteogenic amino acid involved in a wide variety of biological functions, was investigated as an example of the complexity that multiple rotatable bonds produce even in relatively simple molecules. To efficiently explore such a landscape, this molecule was studied in the neutral and zwitterionic forms, in the gas phase and in water solution, by means of molecular dynamics and the enhanced sampling method metadynamics with classical force-fields. Multi-dimensional free energy landscapes were reduced to bi-dimensional maps through the non-linear dimensionality reduction algorithm sketch-map to identify the energetically stable conformers and their interconnection paths. Quantum chemical calculations were then performed on the minimum free energy structures. Our procedure returned the low energy conformations observed experimentally in the gas phase with rotational spectroscopy [M. E. Sanz *et al.*, Phys. Chem. Chem. Phys. **12**, 3573 (2010)]. Moreover, it provided information on higher energy conformers not accessible to experiments and on the conformers in water. The comparison between different force-fields and quantum chemical data highlighted the importance of the underlying potential energy surface to accurately capture energy rankings. The combination of force-field based metadynamics, sketch-map analysis, and quantum chemical calculations was able to produce an exhaustive conformational exploration in a range of significant free energies that complements the experimental data. Similar protocols can be applied to larger peptides with complex conformational landscapes and would greatly benefit from the next generation of accurate force-fields. © *2017 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).* [http://dx.doi.org/10.1063/1.4979519]

I. INTRODUCTION

Aspartic acid (Asp) is a non-essential α-amino acid involved in a number of metabolic functions, from the synthesis of proteins and biomolecules to the gluconeogenesis and the activation of N-methyl-D-aspartate (NMDA) receptors, important for synaptic plasticity and memory functions. Asp side chain is a carboxyl group, usually deprotonated and negatively charged under physiological conditions, which is present in numerous protein active sites.[1] This amino acid is found in the neutral form in the gas phase and in the zwitterionic form in aqueous solution and in crystals. In the latter form, the hydrogen of the Cα–COOH group is donated to the nitrogen to form an NH$_3^+$ amino group, as shown in Fig. 1, resulting into opposite charges for these two moieties.[2] Due to the presence of multiple rotatable bonds, there are many possible (meta)stable conformers for both these forms, which can be exhaustively characterized by the six torsional angles labelled in Fig. 1: three along the backbone, i.e., α (O$_6$–C$_1$–C$_2$–C$_3$), β (C$_1$–C$_2$–C$_3$–C$_4$), and γ (C$_2$–C$_3$–C$_4$–O$_8$); two to

describe the relative position of hydrogen atoms in the carboxyl groups, i.e., δ (C$_2$–C$_1$–O$_6$–H) and ζ (C$_3$–C$_4$–O$_8$–H); and one for the rotation of the amino group, i.e., ϵ (C$_2$–C$_3$–N–H$_N$). ζ is not defined in the zwitterionic form because of the absence of the O$_8$–H hydroxyl, while γ loses relevance due to the symmetry of the remaining carboxylate group. Even for a simple amino acid such as Asp, the flexibility of its torsional angles makes a complete exploration of its conformational free energy landscape far from trivial. In general, the larger is the molecule, the more complex this problem becomes; however, a thorough knowledge of the conformational preferences, and thus of the behaviour, of the building blocks of proteins represents a first crucial step towards the understanding of larger biomolecules. Experimental information on low energy conformers can be obtained by spectroscopic methods (see, e.g., Refs. 3–10). The most powerful technique for conformational identification in the gas phase is rotational spectroscopy, which has been successfully applied in combination with laser ablation[11,12] to determine the conformations of many natural amino acids.[13–16] In the condensed phases, methods such as NMR and X-ray crystallography are also widely used.[17–20] Theoretically, stable/metastable conformers can be identified, for example, by scanning the torsional angles at small intervals,

a)Current address: Department of Chemistry, University College London, Gower Street, London WC1E 6BT, United Kingdom.
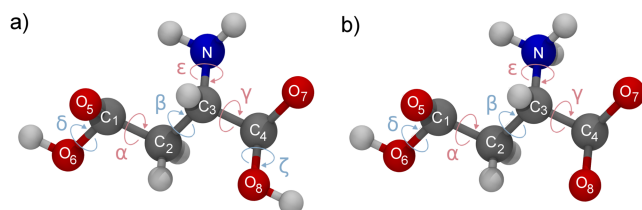
FIG. 1. Neutral (left, (a)) and zwitterionic (right, (b)) aspartic acid, with relevant atoms and torsional angles labelled.

calculating the potential energy for each optimized structure at constrained torsional angles with force-fields or more accurate methods like density functional theory (DFT) and second order Møller-Plesset perturbation theory (MP2), and correcting it *a posteriori* for temperature effects and entropic contributions from normal mode analysis to estimate the free energy.[21,22] Rotational constants and quadrupolar moments can then be calculated to interpret experiments. However, this protocol is cumbersome even for a few torsional angles; it does not calculate the free energy directly at finite temperature, disregarding usually anharmonic effects (although these can be calculated to some extent by second order vibrational perturbation theory at an increased computational cost[23]), and, in aqueous solution, does not consider the solvent explicitly.

Hence, here we suggest an alternative approach to explore the conformational landscape of Asp in a relative wide range of free energies and to benchmark our theoretical protocol. Experimental information (in the gas phase) derived from rotational spectroscopy is available for this amino acid, which has a tractable size, making it ideal for this computational study.[15] The proposed approach is based on the calculation of the free energy landscape at room temperature by means of the enhanced sampling metadynamics technique,[24] as a function of selected torsional angles. The exploration was performed with classical force-fields, to ease the computational cost with respect to *ab initio* calculations, and the results of different parametrisations were compared in order to assess their influence on the free energy. When needed, the dimensionality reduction algorithm sketch-map[25] was employed to simplify the interpretation of the data and help discriminate free energy basins and their interconnecting paths. Once the minimum free energy conformers were identified, their stability and relative energies were evaluated at the DFT and MP2 levels of theory. Besides the low energy conformers in the gas phase accessible to rotational spectroscopy experiments, high energy ones and conformers in solution could also be extracted.

II. METHODS

Preliminary tests showed that the paths separating some of the conformers were too high in free energy to allow transitions between minima in the limited time scale accessible to molecular dynamics (MD), and thus required advanced simulation techniques to accelerate the relevant "rare" events. Metadynamics[24] enhances the sampling of conventional MD by discouraging the system to explore previously visited regions of the configurational space, thus pushing it towards unexplored areas. This is achieved by means of a history-dependent repulsive potential, which biases the dynamics and is built on-the-fly

as a sum of Gaussians deposited along the system trajectory in the space of a subset of carefully chosen slowly varying degrees of freedom, or collective variables (CVs), which describe the process of interest. Once the exploration of the CV space is complete, this history-dependent potential provides a cast of the free energy landscape as a function of the selected CVs.[26–28] Metadynamics has been successfully used to study a wide range of problems in the physical, chemical, biological, and material sciences.[26,29–34]

As metadynamics efficiency decreases with the number of biased CVs, the minimal CV set providing meaningful results should be chosen. To map the conformers of Asp, the three backbone torsional angles α, β, and γ are an obvious, but not adequate choice, as we observed with preliminary calculations. In fact, the free energy as a function of α and γ is characterised by minima separated by barriers of few kcal/mol, which can be adequately sampled even by conventional molecular dynamics. However much higher barriers were observed between the minima as a function of δ and ζ, which define the *trans* and the *cis* isomers of the carboxyl groups. These two angles are involved in the formation of intra-molecular hydrogen bonds, with the amino group acting as a donor or an acceptor when interacting with the hydroxyl and/or the oxygen of the carboxyl groups at the opposite ends. Such interactions can influence the position of the carboxyl groups (e.g., by restraining their mobility when the hydrogen bond is present), thus creating a correlation between the two torsional angles, α and γ, which define the orientation of the groups, and the two torsional angles which describe the position of the carboxylic hydrogens, δ and ζ, respectively. Hence, the most relevant torsional angles driving conformational changes in Asp are β, δ, and ζ, which were used as CVs in the metadynamics simulations. Finally, low free energy barriers, which can be easily sampled by MD, characterize the torsion about ϵ, which determines the orientation of the amino group, justifying its exclusion from the CV selection. For the zwitterionic form, the CVs were limited to β and δ, because of the symmetry of the carboxylate group which makes γ negligible to identify stable conformers, while ζ is not defined.

Asp was simulated with the NAMD 2.9[35] MD package and CHARMM36,[36] one of the most commonly used force-fields in biomolecular simulations, under different conditions: the neutral form in the gas phase, used as reference and for comparison with rotational spectroscopy experiments, and the zwitterionic form in water solution. Two non-physical control systems, i.e., neutral Asp in water and zwitterionic Asp in the gas phase, were also studied to better understand the effect of the environment on the force-field parameters. The two latter cases can be simulated only thanks to the predefined unbreakable bonds in the force-field, which prevent the proton transfer between the amino and carboxylate groups that would naturally occur. To further assess the impact of the choice of the force-field, the calculations for neutral Asp *in vacuo* were repeated with two other popular parametrisations, i.e., AMBER ff99SB-ildn[37] and OPLS-AA,[38,39] and the resulting free energy maps were compared to identify the common features and differences.

Restrained ESP charges were calculated at the Hartree–Fock level of theory with a 6-31G* basis set (for consistency

with the force-field parameters), averaging the charges of the molecular geometries of the six conformers identified experimentally,[15] which were first optimised with MP2 and the 6-311++ G(d,p) basis set, using GAUSSIAN 09.[40]

In the gas phase, the molecule was first equilibrated at 300 K with no periodic boundary conditions, while a 12 Å buffer of TIP3P water molecules was added in the solvated systems in periodically repeated truncated octahedral cells and equilibrated at 300 K and 1 atm. A Langevin thermostat with a collision frequency of 1 ps^{-1} and a Langevin piston with a period of 0.1 ps and a decay time of 0.05 ps were applied. The cutoff for long-range interactions was 10 Å and particle mesh Ewald was employed for the calculation of long range electrostatic interactions. No restraints were applied to the vibration of bonds containing hydrogens. A time step of 0.5 fs was used. 100 ns of well-tempered metadynamics simulations[27] were then performed using the metadynamics plug-in PLUMED 1.3[41] in the NVT ensemble, with the average volume from the equilibrated NPT molecular dynamics for the solvated systems. All other parameters were kept the same as those used in the MD simulations. Gaussians with an initial height of 0.2 kcal/mol and a width of 9° were deposited every 0.25 ps to bias the torsional angles, with a bias factor of 5. Helmholtz free energy landscapes were reconstructed after 400 000 Gaussians were deposited and the CVs showed a diffusive behavior.

In a multi-dimensional free energy landscape, the identification of the stable/metastable conformers and their connection network may be non-trivial, even when the dimensions are limited to three as in this case; projections to two-dimensions by integrating out the third (or any additional) CV may lose significant information. To overcome this limitation (which becomes severe when increasing the number of variables), we adopted sketch-map,[25,42,43] a recent method based on metric multi-dimensional scaling,[44] to disentangle the conformational landscape of Asp by mapping it from the high-dimensional space of the metadynamics CVs and other unbiased torsional angles to a bi-dimensional plane. This dimensionality reduction algorithm has been specifically designed to deal with atomistic simulation data, usually characterised by a set of well separated basins interconnected by transition paths. In both biased and unbiased simulations, the trajectory data are characterized by noise due to thermal fluctuations and to uneven sampling of different regions in conformational space. Sketch-map addresses these problems by focusing on reproducing accurately the relative position of basins that are adjacent to each other. By doing so, it can often generate an intuitive representation of complex, high-dimensional or non-Euclidean networks of conformers, and the transition paths between them are kept contiguous. One should always keep in mind, however, that, in some cases, it may just not be possible to represent the full network in two dimensions and hence the sketch-map projection may contain discontinuities. This is the case when periodicity (e.g., of torsional angles) is involved. While the location of such artifacts depends on the details of the reference landmarks and the optimization procedure, they tend to appear in the poorly sampled transition-state regions, so that the different local minima and their relative free energies are typically represented faithfully. For instance, one may find that basins that are

separated by a high free-energy barrier are projected far away from each other even though they might be contiguous in conformational space. Overall, the possibility of visualizing multiple conformers in two dimensions, avoiding the superposition of distinct configurations, generally outweighs the inconvenience of missing some information on the connectivity of different basins.

The sketch-map algorithm requires first the choice of which degrees of freedom to map; then a limited set of landmark points has to be selected, so that they respect the underlying Boltzmann distribution of the trajectory points in the multi-dimensional space. In our case, 500 landmark points were selected according to this criterion and mapped through a multi-step minimisation of a sigmoid function, which is responsible for the mapping of distances. This function was chosen instead of the conventional Euclidean distance for its efficacy in maintaining the distance relationship between points during the projection to a lower-dimensionality space: points that are far or close together in the initial high-dimensionality space are projected far or close together in the final map, at the cost of losing the quantitative value of their distances. The specific parameters of the sigmoid function have to be chosen so to clearly separate the different conformational basins; the resulting map is not unique.[43] We used sketch-map to reduce to bi-dimensional maps of the free energy landscape as a function, e.g., of the three metadynamics CVs β, δ, and ζ and of α, β, and γ.

Once the data had been mapped and the bi-dimensional free energy network calculated, the characteristic conformations belonging to each basin were extracted: a root mean square deviation (RMSD) based cluster analysis of the structures within the trajectories showed that, once cleaned of noise, a single strongly favoured conformation was present in each basin in most cases. In a few cases, when mapping the three-dimensional space of the metadynamics CVs, more than one conformer was associated with a minimum free energy basin, as information on the α and γ angles had been partially lost because of the selection of the mapped CVs. The cluster analysis helped disentangle the stable degenerate conformers within the same basin: in fact, the free energy dependence on α and γ is strongly influenced by the presence of intramolecular hydrogen bonds, making it easy to extrapolate data on all the six torsional angles and partially recover the information hidden by the CV selection. To verify that we were able to achieve a full discrimination of the conformers, we also produced sketch-maps based on the sampling of up to five torsional angles, e.g., only excluding the fast degree of freedom described by ϵ. We averaged the free energy values of each basin weighted by the corresponding Boltzmann factor, appropriately normalized. We estimated the error by block analysis: after convergence was reached, the data corresponding to the last 2.5 ns of the metadynamics simulation were divided into twelve blocks and populations and free energies were calculated separately for the various blocks. The reported error of the average free energy is the standard deviation for these blocks divided by the square root of their number.

In order to distinguish the identified conformers, we adopted the following nomenclature, loosely following Ref. 15. The free energy profile of the central torsional angle

β is characterised by three minima: $a \simeq \pm 180°$, $b \simeq -60°$, and $c \simeq 60°$. These have the strongest impact on the sketch-map, which is circularly arranged around the values of β. The δ and ζ profiles have two minima related to the *cis* and *trans* configurations of the COOH groups: δ and ζ, $\simeq \pm 180°$ correspond to the *cis* configuration (C), and $\simeq 0°$ to the *trans* (T). We labelled each conformer according to the order of δ, β, and ζ. Moreover, although slightly redundant, we added extra information to account for the intra-molecular hydrogen bonds that the specific conformer forms and to make the different interactions identifiable at a glance. Intra-molecular hydrogen bonds were here counted when the donor-acceptor distance was less than a 3.5 Å cutoff and when the donor-H···acceptor angle was larger than 100°, a criterion previously used for intra-molecular hydrogen bonds.[45] Specifically, we labelled type 1 a hydrogen bond where the amino group acted as a donor to the carbonyl oxygen, as in N–H···O–C, and type 2 a hydrogen bond where the amino group acted as an acceptor and the O–H group as a donor, as in O–H···N. In type 3 hydrogen bonds, the amino group acted as a donor, but the acceptor was the oxygen of the hydroxyl group of COOH, i.e., N–H···O–H. These type numbers were noted to the left of the δ, β, ζ sequence if the orientation of the COOH group involved in the hydrogen bond was described by δ and on the right if it was described by ζ. When the two carboxylic groups were interacting, as in O–H···O, the hydrogen bond was of type 4: in this case, the hydrogen bond label was attributed to the donor to avoid double counting. In the few cases where the same group formed multiple hydrogen bonds, both types of bond were noted.

The stable conformers identified with the CHARMM36 force-field were then optimized both with DFT and the B3LYP exchange and correlation hybrid functional and with MP2, with a 6-311++ G(d,p) basis set. The corresponding free energies were estimated by adding temperature effects and entropic contributions, based on normal modes analysis, using the thermochemistry algorithm in GAUSSIAN09 at ambient conditions.[40] DFT and MP2 calculations in water solution were carried out using a polarizable continuum model. This allowed us to confirm the stability of such conformers and compare energies and ranking obtained with the empirical force-fields to those obtained with quantum chemical calculations. While in principle these calculations provide Gibbs free energies, the PV terms in the enthalpy are calculated within the ideal gas approximation and only depend on temperature. Hence it is identical for different conformers of the same molecule and cancels out when considering relative free energies, making relative Gibbs free energies equal to relative Helmholtz free energies. For a direct comparison with the force-field results, we also employed a similar scheme, as implemented in the GROMACS 5.1.4 simulation package,[46] to calculate the harmonic free energies for the conformers of Asp in the gas phase with the CHARMM36 force-field. This also provided information on anharmonic effects captured with the metadynamics simulations.

In principle, metadynamics simulations can be also performed within quantum chemical schemes that support force calculations and molecular dynamics simulations, e.g., DFT. However, to obtain converged free energies, the simulations need to be run for ~100 ns, as done with the force-field, and become prohibitively expensive, in particular when explicit solvent is also included. Hence we opted for the exploration of the free energy landscape with a force-field, which allowed us to identify the minimal free energy conformers and their interconversion paths, and to complement with single-point quantum chemical calculations as described above.

III. RESULTS AND DISCUSSION

A. Neutral Asp in the gas phase

In Fig. 2, the projections of the CHARMM36 free energy maps for neutral Asp as a function of pairs of torsional angles are shown. Although these projections give a rough idea of the arrangement of basins, details regarding the free energy of individual minima and the complexity of the network of paths and barriers among them are lost in the integration procedure over the third variable. To complement these projected maps and disentangle the network of interconnected basins, a bi-dimensional sketch-map of the free energy landscape as a function of β, δ, and ζ was produced with sigmoid function parameters,[43] σ = 2.0, a_D = 8.0, b_D = 2.0, a_d = 6.0, b_d = 2.0. The result is in Fig. 3. Here, 12 minimum free energy basins are identified and labelled with capital letters in order of increasing free energy; the corresponding structures are also shown.

A typical weakness of non-linear dimensionality reduction methods is the difficulty to reproduce the space in presence of topological obstructions without losing information on the distances between points and introducing artifacts, such as tearing, in the connectivity.[47] However, here the sketch-map
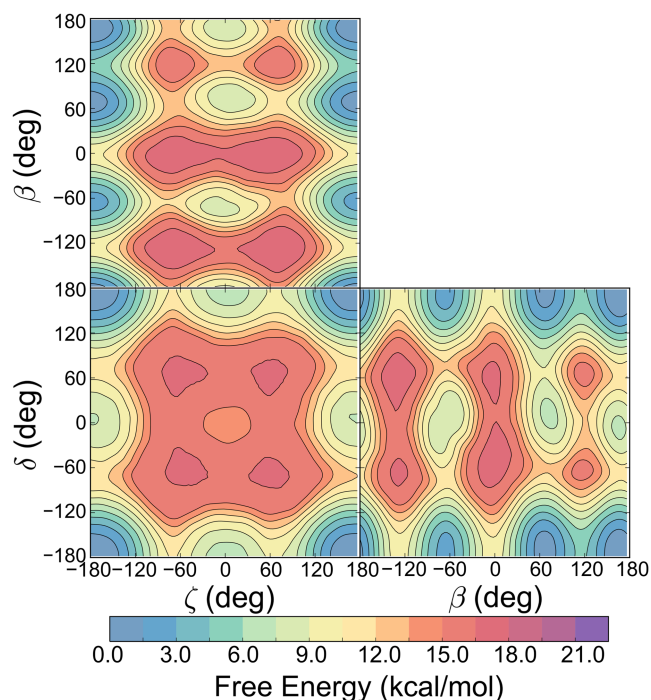

FIG. 2. Projected free energy maps for neutral Asp in the gas phase as a function of pairs of torsional angles calculated with metadynamics and the CHARMM36 force-field. The zero of the free energy scale is set at the absolute minimum.
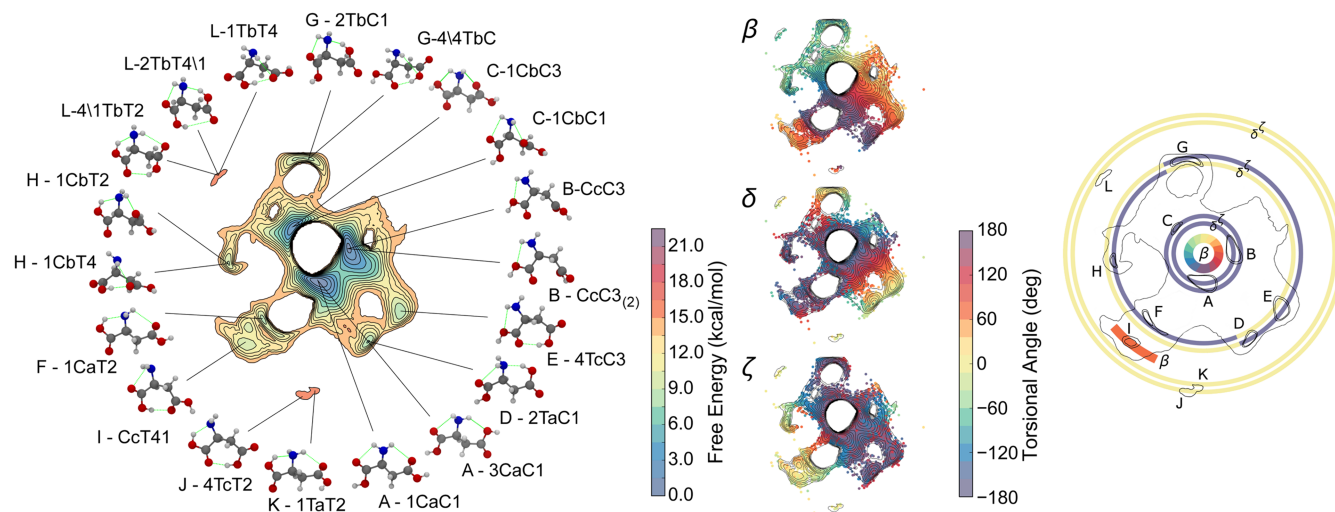
FIG. 3. Left: sketch-map of neutral Asp in the gas phase. The minimum free energy conformers are shown and labelled, with the front letter indicating the name of the basin they belong to; intra-molecular hydrogen bonds are in green. The zero of the free energy scale is set at the absolute minimum. Center: distribution of the values of β, δ, and ζ onto the sketch-map. Right: schematic distributions of the β, δ, and ζ values for the minima onto the sketch-map. β values are displayed in the inner circle, with the exception of the minimum I. δ and ζ values are shown by a pair of circles at different radii corresponding to the minima locations (δ: internal circle and ζ: external circle).

algorithm is remarkably able to preserve the periodicity of the primary torsional angle β and displays the bi-dimensional map circularly around it, showing tearing only in the two other torsional variables and making the interpretation of the data and the association of conformers to the minima intuitive. The map is indeed circular around the value of β, which defines the main backbone variations in the Asp conformers. This is also shown schematically in the right panel of Fig. 3; in the central panel, the angle distributions onto the sketch-map demonstrate that the specific choice of sketch-map parameters is able to separate conformers characterized by different torsional angles. In particular, the bottom region with respect to the central hole is mainly characterised by conformers with $\beta \simeq \pm180°$, the top left region with $\beta \simeq -60°$, and the top right region with $\beta \simeq 60°$. The exception is represented by minima I which has $\beta \simeq 60°$, but is disjointed from the conformers with a similar β value.

Three minima A, B, and C have significant lower free energies than the rest of the map and are characterised by both δ and ζ in the *cis* isomer of the carboxyl group. Transitions between the A and B basins are favoured with respect to transitions to the C basin as can also be inferred from Fig. 2. The *cis* isomers for both carboxyl groups make the formation of type 2 hydrogen bonds impossible, while type 1 hydrogen bonds are frequent and more favourable with respect to type 3 or 4, which are not observed in these basins. In fact, it is a characteristic of type 2 hydrogen bonds to have the COOH group in the *trans* arrangement; the establishment of a hydrogen bond compensates the higher-in-energy COOH *trans* configuration with respect to the *cis*. The conformer with $\beta \simeq \pm180°$, i.e., with a fully extended backbone, is the most favourable and its free energy is set as the zero reference for the calculations made with the CHARMM36 force-field. A second concentrical tier of several higher free energy minima is found around the three main ones. These contain molecules where either δ or ζ have values corresponding to the *trans* arrangement of the respective COOH group, suggesting that *trans* conformers are

overall disfavoured. This is confirmed also by the presence of a third scarcely populated tier, with only three high free energy minima, with both δ and ζ corresponding to conformers with both COOH groups in a *trans* configuration; no connection to other basins appears because of the high free energy and consequently insufficient sampling. In this tier, the free energy is up to 13.2 kcal/mol above the lowest-energy conformer, and few stable conformers can be extracted. The free energy of the minima belonging to the second tier, from D to I, is between 7 and 8 kcal/mol above the lowest-energy conformer and again, type 1 hydrogen bonds are the most frequent, although the presence of hydrogen bonds of types 2, 3, and even 4 is also observed.

From these 12 basins, 19 distinct stable conformers were extracted, the increased number being due to the free energy degeneracy of conformers with different α or γ angles: each structure is shown in Fig. 3, linked to its respective basin. Degeneracy occurs when the presence of intra-molecular hydrogen bonds does not hinder the torsion about α and/or γ. Exemplary is the case of the minimum B, where no intra-molecular hydrogen bonds are present and thus two similar conformers, with complementary orientations of α, are observed. Similarly, rotations about ϵ, which defines the position of the amino group hydrogens, should introduce degeneracy. However, isomers with different values of ϵ, e.g., with hydrogen bond interactions with either one or the other hydrogen of the amine, show little or no difference in terms of free energy. For the sake of clarity, we decided to disregard this extra degree of freedom which gave no useful information.

We also produced a sketch-map which mapped the five torsional angles α, β, γ, δ, and ζ to enforce a full separation of the conformers. However this did not produce additional information with respect to the simpler sketch-maps on β, δ, and ζ and the cluster analysis and is therefore not shown for simplicity; moreover, details on the high energy conformers were lost. It was however an important test to check that our analysis captured all significant conformers.

TABLE I. The relative free energies of the conformers of neutral Asp in the gas phase relative to their absolute minimum calculated with the CHARMM36 force-field and metadynamics simulations (FF-meta) and the harmonic approximation (FF), DFT-B3LYP and MP2.

Basin	Conformer	Free energy (kcal/mol)			
		FF-meta	FF	DFT	MP2
A	1CaC1	0.00 ± 0.01	0.03	0.70	0.86
A	3CaC1	0.00 ± 0.01	0.00	3.00	3.00
B	CcC3	0.73 ± 0.02	0.25	2.65	2.23
B	CcC3$_{(2)}$	0.73 ± 0.02	1.52	4.06	3.43
C	1CbC1	2.11 ± 0.05	3.07	0.76	0.00
C	1CbC3	2.11 ± 0.05	4.37	0.97	3.08
D	2TaC1	6.89 ± 0.06	6.38	0.93	1.21
E	4TcC3	7.24 ± 0.06	4.56	5.95	6.33
F	1CaT2	7.46 ± 0.08	7.48	0.00	0.69
G	2TbC1	7.66 ± 0.13	5.14	2.43	1.98
G	4\3TbC	7.66 ± 0.13	5.83	5.79	6.26
H	1CbT2	7.87 ± 0.15	7.03	0.45	0.36
H	1CbT4	7.87 ± 0.15	5.05	4.50	5.07
I	CcT4\1	8.27 ± 0.10	5.52	2.49	3.34
J	4TcT2	12.17 ± 0.53	11.82	3.96	5.33
K	1TaT2	12.17 ± 0.49	14.66	5.98	6.45
L	4\1TbT2	12.18 ± 0.44	11.65	4.45	4.41
L	2TbT4\1	12.18 ± 0.44	10.94	8.22	7.77
L	1TbT4	12.18 ± 0.44	16.13	9.49	10.17

In Table I, the free energy of the conformers extracted with metadynamics and sketch-map (on β, δ, and ζ) for neutral Asp with respect to their respective absolute minimum is shown. The table includes the average free energies of the basins calculated with metadynamics and the CHARMM36 force-field (FF-meta), and those obtained for the optimized structures by including temperature and entropic effects using normal mode analysis, with CHARMM36 (FF), DFT-B3LYP, and MP2.

Quantum chemical calculations confirm that the identified conformers are all minimum energy structures. However, there are some discrepancies between the free energies obtained with the force-field and those obtained with quantum chemical calculations in the free energy ranking; DFT-B3LYP and MP2 results are overall in good agreement with differences mostly within chemical accuracy (<1 kcal/mol). According to these calculations, 1CbC1 and 1CaT2 found in the basins C and F are the absolute minima for MP2 and DFT-B3LYP, respectively, closely followed by 1CaC1 in basin A. Nevertheless, when making these comparisons, one should bear in mind also the differences between the adopted methodologies; while the force-field metadynamics simulations were carried out at finite temperature and the free energies were calculated directly on the basins, the B3LYP-DFT and MP2 calculations were performed over the minimal potential energy structure at absolute zero, with temperature effects and entropy corrections added *a posteriori*, disregarding, e.g., anharmonic effects. By comparing the free energy differences from metadynamics, those obtained within the harmonic approximation with the CHARMM36 force-field, and the static force-field minimum energies, we can observe that the mean absolute error on the free energy differences due to disregarding anharmonic effects is 1.4 kcal/mol, which is larger than the mean absolute harmonic correction (0.5 kcal/mol) to the free energy

differences. The fact that even for a relatively simple molecule in the gas phase anharmonic corrections are significant underscores the importance of considering them in the analysis of complex biomolecules. Conformer 1TbT4 is a stable high energy conformer, scarcely populated, that was found in basin L together with lower energy conformers.

Experimentally, six conformers were observed with rotational spectroscopy. The results are summarised in the work of Ref. 15, where the estimated relative abundances did not fully correlate with *ab initio* free energy calculations at the MP2 level. These experiments were conducted in supersonic jets. On the initial stages of the supersonic expansion, collisions of the different conformers in the jet with the carrier gas molecules (typically noble gases) can lead to conformational relaxation of higher energy conformers to the lowest energy ones if the interconversion barriers are sufficiently low. For systems involving only one degree of freedom, such as torsional isomerism or axial/equatorial equilibrium in hydrogen bond complexes, barriers less than about 400 cm^{-1} (1.1 kcal/mol) have been proposed to allow relaxation. In multiconformational systems, where several degrees of freedom can channel the relaxation, it has been suggested that barrier heights of 1000 cm^{-1} (2.9 kcal/mol) or higher would impede conformational relaxation. The "effective population" of the observed conformers in the supersonic jet is a consequence of population due to their relative energies plus the possible contributions from relaxation of other conformers.[48,49] Using our nomenclature, the relative abundances estimated in experiments follow the ranking 1CaC1 > 1CbC1 ≈ 2TaC1 > 1CaT2 ≈ 2TbC1 > 1CbC3, with conformers attributed according to the MP2 calculations of the rotational constants.[15] Correlating relative free energies and abundances, this trend is in overall good agreement with the results obtained with metadynamics and the CHARMM26 force-field, with the exception of 1CbC3, which belongs to basin C and is thus lower in free energy than expected. The choice of mapping only the metadynamics CVs makes this conformer degenerate with 1CbC1 from which it differs only for the orientation in α of the carboxyl group.

B. Zwitterionic Asp in water

The free energy surface obtained with metadynamics of the solvated Asp in its zwitterionic form is shown in Fig. 4 (top panel) as a function of β and δ; here, six minima are observed. The symmetry of the COO$^-$ carboxylate moiety allows us to reduce the number of CVs to two (β and δ) without losing crucial information; ζ is non-existent in this system, due to the transfer of the carboxyl oxygen to the nitrogen group. There is in principle no need to use sketch-map in this case because the map is already bi-dimensional; however, in order to clarify the process of deformation involved in the application of the algorithm, a two CV sketch-map, computed with sigmoid parameters, $\sigma = 2.0$, $A = 8.0$, $B = 2.0$, $a = 6.0$, $b = 2.0$, is also shown in Fig. 4. This sketch-map transforms the toroidal surface of the two periodic CVs into a map in the Euclidean plane, with the consequent tearing of connections like in the case of basins D and F.

10 stable conformers, which account for the free energy degeneracy with respect to the α torsion, were extracted with a structural cluster analysis and associated with their respective
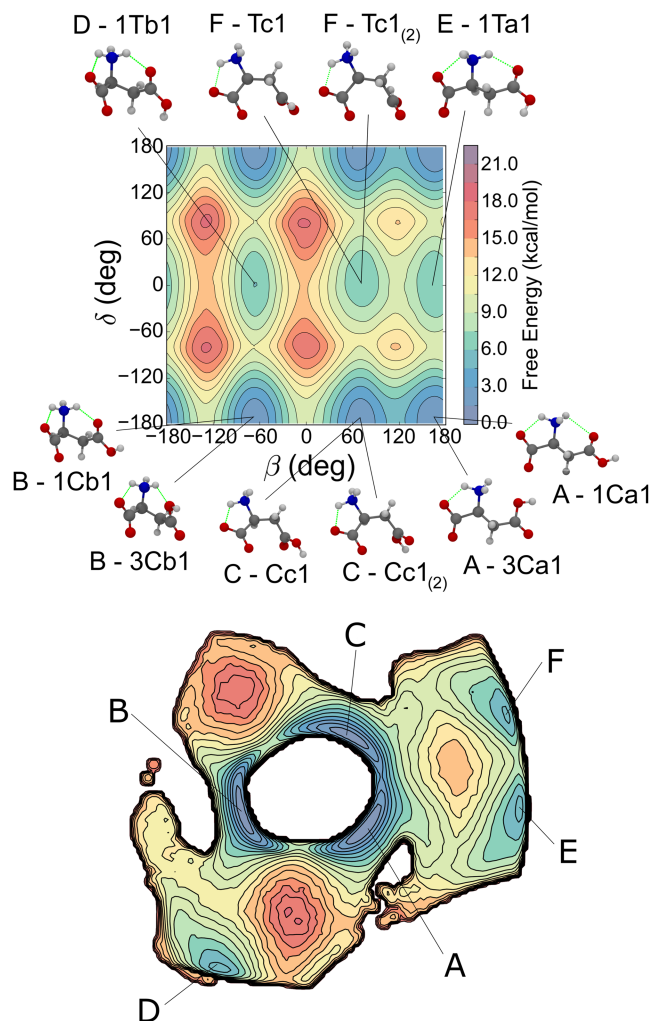
D - 1Tb1   F - Tc1   F - Tc1$_{(2)}$ E - 1Ta1

B - 1Cb1

B - 3Cb1   C - Cc1   C - Cc1$_{(2)}$ A - 3Ca1

A - 1Ca1

FIG. 4. Top: free energy surface of zwitterionic aspartic acid in water as a function of β and δ. The zero of the free energy scale is set at the absolute minimum. The conformers extracted are shown and labelled with the front letter indicating the basin they belong to. Intra-molecular hydrogen bonds are shown in green. Bottom: bi-dimensional representation of the free energy obtained with sketch-map.

minima. We verified that this list exhausts all the favourable isomers of zwitterionic Asp. Hence, an additional sketch-map starting from the four-dimensional space of the α, β, γ, and δ angles to eliminate degeneracies was not necessary. The only two remaining conformers, obtained from the rotation of the α angle in basins D and E, require the carboxyl hydrogen to move closer to the positive amino group, in highly unfavourable conformations. In all the stable isomers, a hydrogen bond between the positively charged NH_3^+ moiety and the negatively charged COO^- group is present. This interaction is stabilized in the solvent and does not prevent the formation of further hydrogen bonds with the surrounding water molecules. When comparing the metadynamics map in Fig. 4 with the projection of the three-dimensional free energy landscape of neutral Asp in the gas phase onto the β-δ plane in Fig. 2 (right), interesting similarities emerge. The main change seems to be the decreased free energy difference between *trans* and *cis* conformers in solution with respect to the gas phase.

The free energy has two regions corresponding to the *trans* and *cis* values of δ and, as in the gas phase, the *cis* conformers seem favoured. In the three lowest free energy basins A, B, and

C, the carboxyl group is indeed in the *cis* conformers, while β has values of $\pm 180°$, $\simeq -60°$, and $60°$, respectively. All three minima have similar free energies (within 0.5 kcal/mol) and again are arranged circularly in the sketch-map. Basin A contains conformers with $\beta \simeq \pm 180°$ with the possibility to form a type 1 or a type 3 hydrogen bond involving the COOH group. Similarly, both B and C comprised degenerate conformers characterised by the two possible orientations of α. Type 3 hydrogen bonds were observed in B, while in C no hydrogen bond interaction between the amino group and the carboxylic group can occur. This similarity in the free energies of the A, B, and C basins may be facilitated by the presence of water molecules, which provide a favourable environment for Asp through numerous hydrogen bond interactions with both the carboxyl groups and the amine. For the *trans* conformers in δ (D and E), the double degeneracy in α is lost as orienting the OH group near the NH_3^+ moiety would draw the hydrogens of both groups too close: the repulsion of their positive charges would prevail making these conformers unstable and only type 1 hydrogen bonds were thus observed. This is not the case for basin F, whereas for basin C, the configuration of β prevents the NH_3^+ and the COOH groups to interact; hence, the double-degeneracy with respect to the α-torsion is preserved.

In Table II, the free energies for zwitterionic Asp in a water solution are shown for the 10 distinct conformers obtained from metadynamics and *ab initio* calculations with implicit solvent. The relative free energies calculated with the different methods are more similar in water solution than *in vacuo*. This may be due to the specific force-field parametrization, which is primarily developed for use within a solvated biomolecular environment rather than in the gas phase, and may be therefore more appropriate in this case.

In Fig. 5, the free energy of both the zwitterionic system in water and the neutral system in the gas phase is mapped via sketch-map starting from the three-dimensional space of α, β, and γ which is common to both systems. While the absence of ζ in the zwitterionic simulations prevented a direct comparison with the previous sketch-maps, working with these torsional angles allows us to observe how the free energy is modified by the presence of water molecules. Neglecting δ and

TABLE II. The relative free energies for the stable conformers of zwitterionic Asp in water, calculated with metadynamics and the CHARMM36 force-field (FF-meta), DFT-B3LYP, and MP2 (with implicit solvent), with respect to their respective absolute minimum.

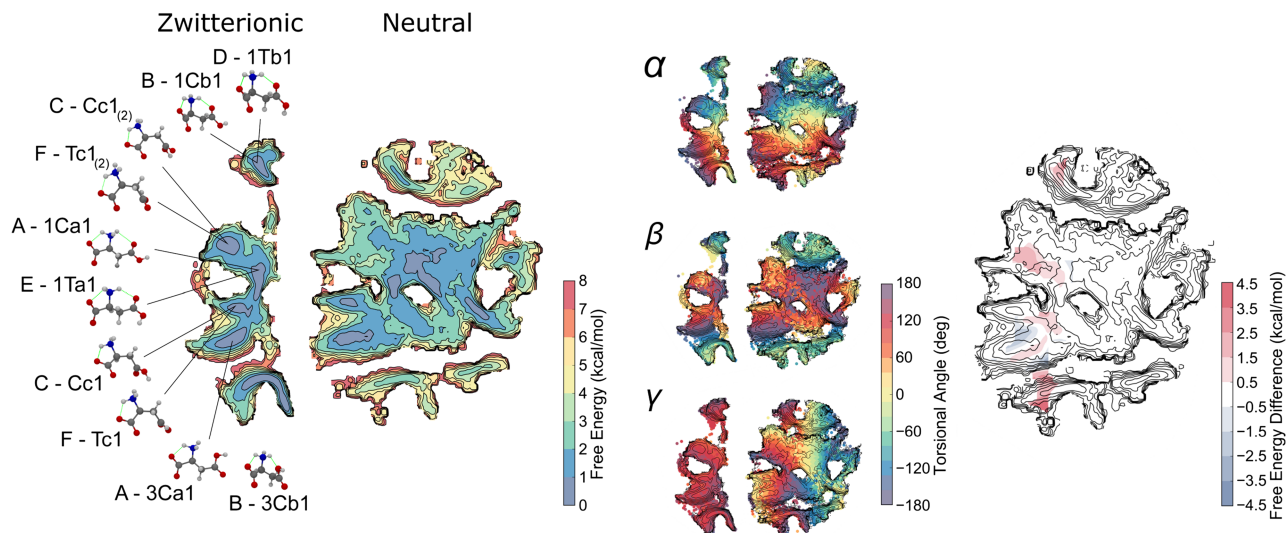| Basin	Conformer	Free energy (kcal/mol)		
		FF-meta	DFT	MP2
A	3Ca1	0.00 ± 0.08	0.21	0.93
A	1Ca1	0.00 ± 0.08	1.48	1.83
B	1Cb1	0.42 ± 0.08	0.00	0.00
B	3Cb1	0.42 ± 0.08	1.62	1.19
C	Cc1	0.47 ± 0.08	2.05	2.56
C	Cc1$_{(2)}$	0.47 ± 0.08	2.38	2.83
D	1Tb1	4.49 ± 0.11	2.53	2.29
E	1Tc1	4.62 ± 0.10	4.45	5.35
F	Ta1	4.86 ± 0.10	3.14	3.61
F	Ta1$_{(2)}$	4.86 ± 0.10	3.14	3.61

FIG. 5. Left: sketch-map of zwitterionic Asp in water and its neutral version in the gas phase obtained by mapping the three-dimensional space of the torsional angles α, β, and γ. The zero of the free energy scale is set at the absolute minimum. The minimum free energy conformers of the zwitterionic form are shown and labelled to help the comparison; intra-molecular hydrogen bonds are in green, while the colour range has been reduced with respect to previous figures to improve the readability. Center: distribution of the values of α, β, and γ onto the sketch-maps. Right: difference between the sketch-map of neutral Asp *in vacuo* and that of zwitterionic Asp in water.

ζ, whose isomerisation heavily affects the free energy landscape, however, causes a non-negligible loss of information, as most basins are now hidden by lower energy minima previously labelled as A, B, and C (for both systems). This makes it more difficult to distinguish quantitatively different conformations and required to decrease the range of free energy in the graph; hence, these maps should be considered complementary to the others.

The most evident effect is the absence of half of the map in the zwitterionic simulation with respect to the vertical axis. This is due to the formation of a stable hydrogen bond between the amino group and the negative carboxyl group. γ, the angle defining the rotation of such group, hence is bound to explore only values around 120° since it was not biased in the metadynamics. The symmetry of this moiety, however, guarantees that the second half of the map would be identically mirrored with respect to the vertical axis. This is not the case for the neutral map, where differences between the two halves arise from the presence of the hydroxyl hydrogen.

While γ is mapped "horizontally," α is mapped "vertically," with conformations with negative α values mapped onto the top half and conformations with positive α values mapped onto the lower half. The basins belonging to different values of β are also easily discernible.

Another interesting feature is the lowering in free energy of certain regions of the map for zwitterionic Asp in water with respect to its neutral counterpart *in vacuo*: specifically, at the top, a basin containing 1Cb1 and 1Tb1, belonging to basins B and D in the previous sketch-map of Fig. 4, in the middle, the basin containing $Cc1_{(2)}$ and $Ta1_{(2)}$, previously in C and F, respectively, and finally at the bottom of the map, a basin containing 3Cb1, previously in B. It is safe to assume that the lowering of such minima is due to the presence of 1Cb1, $Cc1_{(2)}$, and 3Cb1, rather than the other three conformers. In this map, all minima are quite close in free energy, as

already observed in the analysis of the bi-dimensional map of zwitterionic Asp, which is likely to be due to the numerous favourable hydrogen bonds that the conformers can form with the solvent in competition with intra-molecular hydrogen bonds.

C. Unphysical systems: Neutral Asp in water and zwitterionic Asp in the gas phase

The pre-set unbreakable bonds of the force-field allow us to study extremely unfavourable environments: two control simulations of unphysical systems, neutral Asp in water and zwitterionic Asp in the gas phase, have been thus carried out (using the same simulation protocol as before), to analyse the behaviour of this amino acid from a different perspective. It is interesting to analyze the behaviour of Asp in such extreme conditions as they may provide information on the transformation between the neutral and the zwitterionic forms.

The sketch-map for solvated neutral Asp can be seen at the top of Fig. 6 and, when compared with the sketch-map obtained in the gas phase, it allows us to understand the effect of the water environment. The positions of the minima are consistent with those of the gas phase and so is their ranking in most cases, but the relative free energy differences (calculated with the CHARMM36 force-field) between *trans* and *cis* conformers (in δ and ζ) are smaller. Indeed, the free energy difference of the *cis-trans* configurations (in δ or ζ) is about 4 kcal/mol, with respect to 7.5 kcal/mol *in vacuo*, in line with the effect of water on the *trans* conformers observed in Sec. III B for zwitterionic Asp in solution. This is likely to be due to the occurrence of hydrogen bond interactions between the carboxyl groups and the water molecules, which weaken the intra-molecular hydrogen bond network, previously seen to define the morphology of this map in the gas phase, and diminish its effect on the Asp free energy. As a consequence, new minima of relatively high
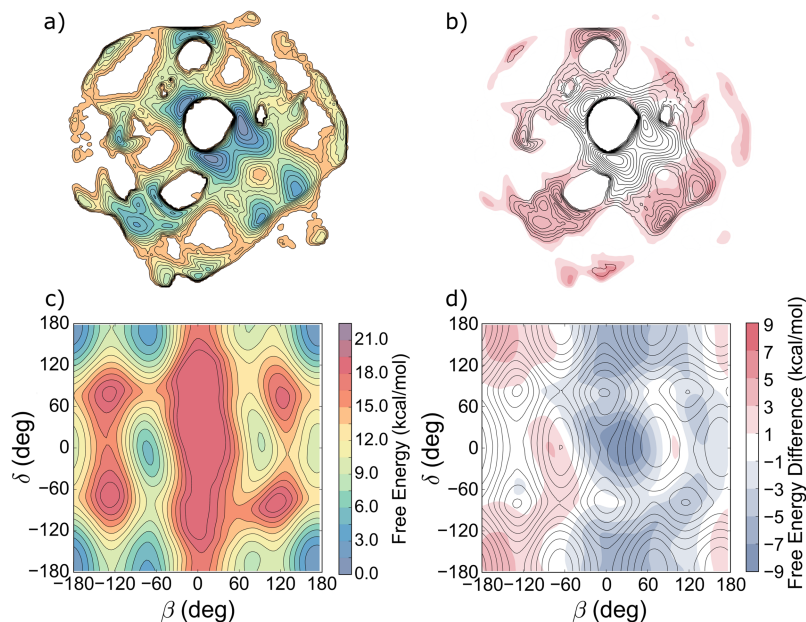
FIG. 6. Top: (a) sketch-map of neutral Asp in water and (b) its comparison with the corresponding sketch-map in the gas phase $\Delta F_{gas-phase} - \Delta F_{water}$. Bottom: (c) free energy map as a function of β and δ of zwitterionic Asp in the gas phase and (d) its comparison with the corresponding map in water $\Delta F_{water} - \Delta F_{gas-phase}$. The colour range goes from blue, where the original physical systems have lower free energy, to white and red, where the physical systems have a higher free energy. The zero of the free energy scales is set at the corresponding absolute minimum.

free energy, which were not visited in the gas phase scenario, are sampled at the bottom and at the periphery of the sketch-map.

The transition paths connecting the basins of the second tier are also lowered by this effect, allowing for easier changes of the values of the δ and ζ angles. *Cis-trans* isomerisation of the two COOH groups is thus more favourable than *in vacuo* with isomerisation barriers of ~6 kcal/mol. The *trans* configuration of the side chain COOH determined by δ is expected to facilitate a direct proton transfer to the nitrogen group during the formation of the zwitterionic molecule. One could sensibly speculate that this is the favourable proton transfer mechanism, although here conformers characterised by direct intra-molecular hydrogen bonds between the amino moiety, acting as acceptor, and the O_8–H hydroxyl group, as donor, are the minima with the highest energy. An alternative mechanism may include the intervention of chains of water molecules as proton carriers: in this scheme, the carboxyl group may donate the hydrogen to the solvent, which could then transport and give it to the amine.

In the map for zwitterionic Asp *in vacuo* at the bottom of Fig. 6, differences with respect to the case in water can be clearly seen. Although the location of the minimum free energy basins remains the same and the absolute minimum still corresponds to basin A, the basins are deformed and there is a change in their relative free energy and ranking. While B was before the second deepest basin, it is here the fourth, with C and D which become instead the second and third lowest free energy minima. In particular, the conformers with $\beta \simeq 60°$ are destabilized with respect to the water solution. The map loses symmetry because of the presence of a wide barrier at $\beta \simeq 0°$ affecting especially the basins at $\delta \simeq 0°$, which are now narrower but higher in free energy for $\beta \simeq 60°$ and 180°. Conformational changes involving β torsions between $\simeq 60°$ and $\simeq -60°$ are hindered by this barrier. All the minimum free energy conformers present a direct interaction between the COO⁻ group and the NH₃⁺ moiety, which would lead, in unconstrained conditions, to a direct proton transfer

without barrier, as easily demonstrated with quantum chemical methods.[50]

D. Force-field assessment

As already observed, in the sketch-map of neutral Asp in the gas phase in Fig. 3, it is notable the arrangement of basins in tiers, which reflect an increase of about 6-7 kcal/mol per tier in the free energies of Table I.

From these data, one could argue that the CHARMM36 force-field penalizes the *trans* configuration of the carboxyl COOH groups with respect to the *cis* by such an amount, increasing the free energy of the basin by $\simeq 7$ kcal/mol whenever one of the two carboxyl groups assumes the *trans* configuration or by $\simeq 12$ kcal/mol when both do, as in the case of the external tier. Since this force-field parametrization is optimized for solvated proteins rather than gas phase simulations, this penalty may arise from its use *in vacuo*, where the absence of networks of hydrogen bonds with the surrounding water molecules may affect the simulation

TABLE III. The relative free energies of neutral Asp conformers in the gas phase with respect to their absolute minimum calculated with metadynamics and the CHARMM36, AMBER99-ildn, and OPLS-AA force-fields.

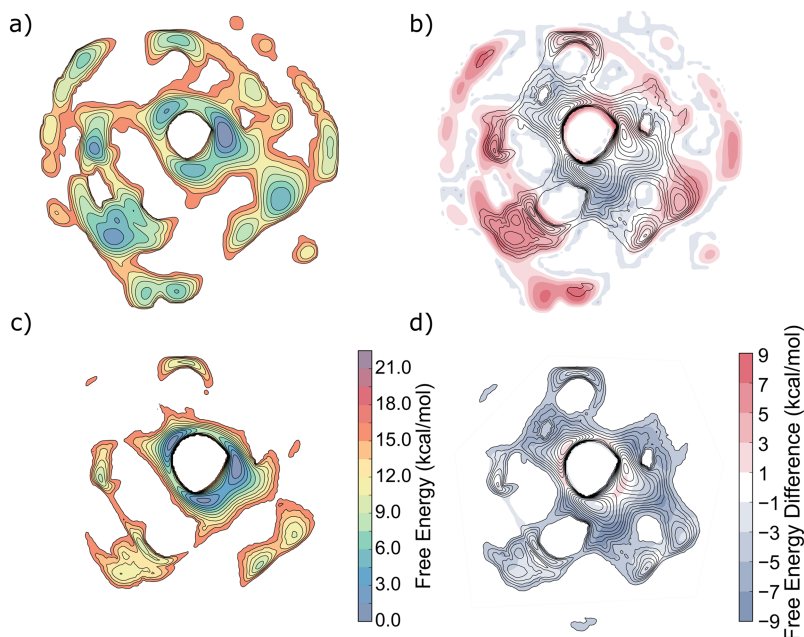| Basin	Free energy (kcal/mol)		
	CHARMM	AMBER	OPLS
A	0.00 ± 0.01	3.35 ± 0.03	1.40 ± 0.02
B	0.73 ± 0.02	0.00 ± 0.03	0.00 ± 0.03
C	2.11 ± 0.05	2.96 ± 0.06	0.63 ± 0.04
D	6.89 ± 0.06	6.30 ± 0.08	8.56 ± 0.07
E	7.24 ± 0.06	3.01 ± 0.17	10.36 ± 0.07
F	7.46 ± 0.08	3.59 ± 0.14	9.63 ± 0.08
G	7.66 ± 0.13	5.31 ± 0.81	10.04 ± 0.13
H	7.87 ± 0.15	2.39 ± 0.57	8.61 ± 0.31
I	8.27 ± 0.10	2.08 ± 0.22	10.49 ± 0.10
J	12.17 ± 0.53	4.46 ± 0.78	
K	12.17 ± 0.49	4.35 ± 0.36	
L	12.18 ± 0.44	5.22 ± 0.57	16.29 ± 0.31

FIG. 7. Left: free energy sketch-maps for neutral Asp in the gas phase calculated with the AMBER ff99SB-ildn (top) and OPLS-AA (bottom) force-fields. The zero of the free energy scales is set at the corresponding absolute minimum. Right: the difference in free energy of the sketch-map obtained with CHARMM36 with respect to those obtained with AMBER ff99SB-ildn and OPLS-AA, respectively.

and could be mitigated with the use of a specific force-field parametrization for the gas phase. The corresponding CHARMM36 results for the zwitterionic molecule *in vacuo* seems to confirm this, as the difference between basins with the δ CV in *cis* or *trans* configuration is decreased to about 4 kcal/mol.

As an assessment of the selected force-field, neutral Asp *in vacuo* was simulated with two other popular force-fields, AMBER ff99SB-ildn and OPLS-AA. In spite of the differences in the relative free energies (Table III), all three force-fields were consistent in maintaining the positions of the principal minima and maxima. This can be observed in Fig. 7, where the sketch-maps obtained with AMBER ff99SB-ildn and OPLS-AA are compared to that with CHARMM36 using the same free energy scale as in previous figures. It is evident that the energy penalty of the *trans* with respect to the *cis* configuration is decreased by AMBER ff99SB-ildn, but it is sensibly increased by OPLS-AA. The penalty for the isomerization of the carboxylic hydrogens is reduced consistently, similarly to what was observed in the case of solvated systems with CHARMM36, which modifies the ranking of the extracted conformers. The previously lowest free energy minimum A has its free energy increased to 3.4 kcal/mol, which is higher than those of minima like C and I, thus making B the new absolute minimum. As before, minima belonging to the outer tier are the least favoured, but with a free energy now lowered to about 4.5 kcal/mol. On the opposite side of the spectrum, OPLS-AA may downplay the role of hydrogen bonds over the penalties of *cis-trans* isomerisation. This force-field reproduces the lowest *cis* conformers similarly to the others, but its sampling of the higher free energy *trans* basins is too sparse due to their increased energy. The sketch-map free energies are calculated from histograms of the conformers population in each region, and therefore the high energy conformers, which are scarcely sampled, do not appear in the map.

In conclusion, the choice of the force-field did not affect the identification of the structures of the stable minimum free energy conformers, but only their relative free energies. CHARMM36 produced results for Asp in the gas phase in between the other two tested force-fields and was able to reproduce reasonably well the estimated abundances of the experimentally identified conformers; results can then be refined with quantum chemical methods. Hence we can consider it an adequate choice, at least at the qualitative level.

IV. CONCLUSIONS

A thorough investigation of the conformational properties of biological systems can be cumbersome even for relatively small molecules with multiple rotatable bonds. Rotational spectroscopy experiments are able to extract information for the most populated low energy conformers in the gas phase and they are usually aided by computational methods to map the potential energy surface and predict conformational abundances.

In this work, we have shown how a powerful protocol, which combines classical MD simulations enhanced by metadynamics and the dimensionality reduction algorithm sketch-map, allowed us to efficiently explore the conformational free energy landscape, at finite temperature and at a relatively low computational cost, of a typical amino acid, Asp, in the neutral and zwitterionic forms, in the gas phase and in aqueous solution. These methods allowed us to recover conformers that were identified with rotational spectroscopy experiments, together with a number of structures that were not observed before because of their higher free energy. The use of metadynamics to efficiently enhance the sampling of the conformational space with respect to conventional MD helped us overcome energy barriers, accessing stable conformers within a relatively wide range of free energies. The dimensionality reduction algorithm sketch-map proved useful in simplifying the representation of the multidimensional free energy, disentangling basins, and their network of interconversion paths.

The use of a classical force-field, which is imposed by the computational cost, has some limitations, especially in the gas phase when using force-fields developed for the wet biomolecular environment. Qualitative consistency, however, was found among the three different force-fields here tested, which produced the same conformers albeit with discrepancies in the relative free energies.

Differences in the energy ranking were also observed with the quantum chemical data, more prominently in the gas phase than in solution. It should however be noted that these calculations are also approximate: in fact, they consist of single-point potential energy calculations corrected *a posteriori* and, at variance from direct free energy calculations, do not include anharmonic entropic effects, which may be significant.

Our protocol for mapping amino acid and peptide conformers at finite temperature can be potentially applied to other more complex cases and can be improved with the use of novel and accurate force-fields, e.g., based on machine learning on high quality training sets, presently under development. These studies would ultimately contribute to the understanding of the conformational complexity of the building blocks of biological matter, which is important in a wide variety of problems and fields, from drug design based on molecular recognition mechanisms to the identification of organic compounds and amino acids in the interstellar environment.

ACKNOWLEDGMENTS

[1] D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 6th ed. (W. H. Freeman, 2013).

[2] K. C. Timberlake, *General, Organic, and Biological Chemistry: Structures of Life* (Pearson Benjamin Cummings, 2006), p. 896.

[3] C. S. Barry, E. J. Cocinero, P. Çarçabal, D. P. Gamblin, E. C. Stanca-Kaposta, S. M. Remmert, M. C. Fernández-Alonso, S. Rudić, J. P. Simons, and B. G. Davis, J. Am. Chem. Soc. **135**, 16895 (2013).

[4] R. Weinkauf, J.-P. Schermann, M. de Vries, and K. Kleinermanns, Eur. Phys. J. D **20**, 309 (2002).

[5] I. Peña, E. J. Cocinero, C. Cabezas, A. Lesarri, S. Mata, P. Écija, A. M. Daly, A. Cimas, C. Bermúdez, F. J. Basterretxea, S. Blanco, J. A. Fernández, J. C. López, F. Castaño, and J. L. Alonso, Angew. Chem., Int. Ed. **52**, 11840 (2013).

[6] J. L. Alonso, V. Vaquero, I. Peña, J. C. López, S. Mata, and W. Caminati, Angew. Chem. **125**, 2387 (2013).

[7] C. Calabrese, Q. Gou, A. Maris, S. Melandri, and W. Caminati, J. Phys. Chem. B **120**, 6587 (2016).

[8] S. Krasnokutski, A. Ivanov, V. Izvekov, G. Sheina, and Y. Blagoi, J. Mol. Struct. **482–483**, 249 (1999).

[9] N. L. Burke, A. F. DeBlase, J. G. Redwine, J. R. Hopkins, S. A. McLuckey, and T. S. Zwier, J. Am. Chem. Soc. **138**, 2849 (2016).

[10] S. Habka, V. Brenner, M. Mons, and E. Gloaguen, J. Phys. Chem. Lett. **7**, 1192 (2016).

[11] S. L. Stephens, D. P. Tew, N. R. Walker, and A. C. Legon, Phys. Chem. Chem. Phys. **18**, 18971 (2016).

[12] A. Lesarri, S. Mata, J. C. López, and J. L. Alonso, Rev. Sci. Instrum. **74**, 4799 (2003).

[13] C. Bermúdez, S. Mata, C. Cabezas, and J. L. Alonso, Angew. Chem., Int. Ed. **53**, 11015 (2014).

[14] J. L. Alonso, C. Perez, M. Eugenia Sanz, J. C. Lopez, and S. Blanco, Phys. Chem. Chem. Phys. **11**, 617 (2009).

[15] M. E. Sanz, J. C. López, and J. L. Alonso, Phys. Chem. Chem. Phys. **12**, 3573 (2010).

[16] I. Peña, M. E. Sanz, J. C. López, and J. L. Alonso, J. Am. Chem. Soc. **134**, 2305 (2012).

[17] T. Kimura, N. Matubayasi, H. Sato, F. Hirata, and M. Nakahara, J. Phys. Chem. B **106**, 12336 (2002).

[18] T. Kimura, N. Matubayasi, and M. Nakahara, Biophys. J. **86**, 1124 (2004).

[19] Y. Wang, D. Wilson, and G. S. Harbison, Cryst. Growth Des. **16**, 625 (2016).

[20] S. T. Rao, Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. **29**, 1718 (1973).

[21] D. A. McQuarrie and J. D. Simon, *Molecular Thermodynamics* (University Science Books, 1999), p. 656.

[22] M. Chen and Z. Lin, J. Chem. Phys. **127**, 154314 (2007).

[23] V. Barone, J. Chem. Phys. **120**, 3059 (2004).

[24] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. U. S. A. **99**, 12562 (2002).

[25] M. Ceriotti, G. A. Tribello, and M. Parrinello, Proc. Natl. Acad. Sci. U. S. A. **108**, 13023 (2011).

[26] A. Laio and F. L. Gervasio, Rep. Prog. Phys. **71**, 126601 (2008).

[27] A. Barducci, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **100**, 020603 (2008).

[28] A. Barducci, M. Bonomi, and M. Parrinello, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **1**, 826 (2011).

[29] B. Ensing, M. De Vivo, Z. Liu, P. Moore, and M. L. Klein, Acc. Chem. Res. **39**, 73 (2006).

[30] C. Melis, G. Bussi, S. C. R. Lummis, and C. Molteni, J. Phys. Chem. B **113**, 12148 (2009).

[31] V. Leone, G. Lattanzi, C. Molteni, and P. Carloni, PLoS Comput. Biol. **5**, e1000309 (2009).

[32] F. Comitani, N. Cohen, J. Ashby, D. Botten, S. C. R. Lummis, and C. Molteni, J. Comput.-Aided Mol. Des. **28**, 35 (2014).

[33] F. Comitani, C. Melis, and C. Molteni, Biochem. Soc. Trans. **43**, 151 (2015).

[34] F. Comitani, V. Limongelli, and C. Molteni, J. Chem. Theory Comput. **12**, 3398 (2016).

[35] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, J. Comput. Chem. **26**, 1781 (2005).

[36] R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, and A. D. Mackerell, J. Chem. Theory Comput. **8**, 3257 (2012).

[37] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, Proteins: Struct., Funct., Bioinf. **78**, 1950 (2010).

[38] W. L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc. **110**, 1657 (1988).

[39] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225 (1996).

[40] G. M. J. Frisch, W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, and J. L. Sonnenberg, GAUSSIAN 09, Revision C.01, Gaussian, Inc., 2009.

[41] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello, Comput. Phys. Commun. **180**, 1961 (2009).

[42] G. A. Tribello, M. Ceriotti, and M. Parrinello, Proc. Natl. Acad. Sci. U. S. A. **109**, 5196 (2012).

[43] M. Ceriotti, G. A. Tribello, and M. Parrinello, J. Chem. Theory Comput. **9**, 1521 (2013).

[44] I. Borg and P. J. F. Groenen, *Modern Multidimensional Scaling: Theory and Applications* (Springer Science & Business Media, 2005), p. 614.

[45] D. Botten, G. Fugallo, F. Fraternali, and C. Molteni, J. Phys. Chem. B **119**, 12860 (2015).

[46] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, SoftwareX **1**, 19 (2015).

[47] B. Hashemian and M. Arroyo, J. Chem. Phys. **142**, 044102 (2015).

[48] P. D. Godfrey and R. D. Brown, J. Am. Chem. Soc. **120**, 10724 (1998).

[49] G. M. Florio, R. A. Christie, K. D. Jordan, and T. S. Zwier, J. Am. Chem. Soc. **124**, 10236 (2002).

[50] A. T. Paxton and J. B. Harper, Mol. Phys. **102**, 953 (2004).