PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2018, **20**, 20712

Experimental and theoretical investigations of infrared multiple photon dissociation spectra of arginine complexes with Zn^{2+} and Cd^{2+}†

Aaron M. Chalifoux,[a] Georgia C. Boles,[a] Giel Berden,[b] Jos Oomens[bc] and P. B. Armentrout*[a]

Received 1st June 2018,
Accepted 23rd July 2018

DOI: 10.1039/c8cp03484b

rsc.li/pccp

Arginine (Arg) complexes with Zn^{2+} and Cd^{2+} were examined by infrared multiple photon dissociation (IRMPD) action spectroscopy using light from a free electron laser. Electrospray ionization generated complexes of deprotonated Arg with Zn^{2+}, $[Zn(Arg-H)]^+$, and Arg with $CdCl^+$, $CdCl^+(Arg)$. Possible low-energy conformers of these species were found using quantum chemical calculations, and their calculated IR spectra were compared to experimentally measured IRMPD spectra. Calculations were performed at the B3LYP/6-311+G(d,p) level for Zn^{2+} complexes and B3LYP/def2-TZVP with an SDD effective core potential on cadmium for $CdCl^+$ complexes. $[Zn(Arg-H)]^+$ was found to adopt a charge-solvated, tridentate $[N,CO^-,N^{\omega\prime}]$ structure where Zn^{2+} binds to the backbone amine, carbonyl oxygen, and side-chain terminal guanidine nitrogen ($N^{\omega\prime}$). The $CdCl^+(Arg)$ species was suggested to be a mixture of a dominant (~85%) charge-solvated, tridentate $[N,CO,N^{\omega\prime}]$ structure where the $CdCl^+$ binds to the backbone amine, carbonyl, and side-chain imine ($N^{\omega\prime}$) and a minor (~15%) bidentate $[N,CO^-](N^{\omega\prime}H_2^+)$ zwitterionic structure where the metal center binds to the backbone amine and carbonyl oxygen with intramolecular proton migration from the hydroxyl to the $N^{\omega\prime}$ guanidine nitrogen (as designated in parenthesis).

Introduction

Zinc finger proteins, found in about 3% of the genes in the human genome,[1] play a significant role in important biological processes such as DNA and RNA recognition and binding.[1,2] Specifically, the classical Cys_2His_2 zinc finger domain (where the central Zn^{2+} ion is ligated by two cysteine and two histidine residues that bind *via* their deprotonated side-chain sulfur and imidazole nitrogen sites, respectively) plays an integral role in DNA binding and is relatively common in mammalian species.[3,4] However, few quantitative, systematic studies have been completed on the effects of mutation on the structural and DNA binding of such domains.[5,6] One particular study evaluated the DNA binding of $Cys_2HisXxx$ domains, where Xxx = aspartic acid (Asp), glutamic acid (Glu), asparagine (Asn), glutamine (Gln), alanine (Ala), and arginine (Arg). Here, the latter two residues were used as negative controls, as no significant Zn^{2+}–Ala and Zn^{2+}–Arg interactions were expected.[5] Interestingly, the His→Ala mutation had only a small effect (about a 10-fold decrease) on DNA binding compared to the analogous wild-type protein. (No comparative quantitative study on Arg mutation was reported.) It was postulated that these zinc fingers could potentially remain folded in the presence of a mutant even if no Zn^{2+} binding with the mutant was observed, thus maintaining the ability to bind to DNA. Notably, structural characterization of the $Cys_2HisXxx$ domains were outside the scope of study, such that the specific interactions of Zn^{2+}–Arg binding (if present) within the context of zinc finger domains remains unknown. Because of the importance of zinc finger domains, the characterization of zinc binding to Cys, His, and other potential mutant residues is particularly important. Thus, the current work explores prototypical Zn^{2+}–Arg binding, where comparison between binding at the basic side chains of Arg and His may allow for the determination of the key interactions that favor His ligation in common zinc finger domains.

In previous work, Gutowski and co-workers conducted *ab initio* calculations to find and characterize low-energy conformers of gaseous Arg and reported that the lowest energy conformers overwhelmingly adopted a canonical form.[7] Upon complexation of Arg with alkali metals, infrared multiple photon dissociation (IRMPD) spectroscopy studies by Williams

[a] *Department of Chemistry, University of Utah, 315 South 1400 East, Room 2020, Salt Lake City, Utah 84112, USA. E-mail: armentrout@chem.utah.edu*
[b] *Radboud University, FELIX Laboratory, Institute for Molecules and Materials, Toernooiveld 7c, NL-6525 ED Nijmegen, The Netherlands*
[c] *van't Hoff Institute for Molecular Sciences, University of Amsterdam, Science Park 904, NL-1098 XH Amsterdam, The Netherlands*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8cp03484b

and co-workers found that both zwitterionic and non-zwitterionic structures were formed, creating salt bridge (SB) and charge solvated (CS) structures, respectively.[8] CS structures (adopted by the smaller lithium and sodium cations) involve binding the metal center to the backbone carbonyl, backbone amine nitrogen, and guanidine imine nitrogen, $N^{\omega\prime}$. Binding of the larger alkali metal cations (potassium, rubidium, and cesium) with Arg forms zwitterionic salt bridges.[8,9] In these structures, the carboxylic acid has transferred its proton to the guanidine side chain while maintaining a hydrogen bond, and the metal cation binds to both oxygens of the anionic carboxylate. Similar observations are anticipated for transition metals like zinc and its congener, cadmium, although differences in binding attributable to the increased ionic radius of the metal ion are possible, as the ionic radius of Zn^{2+} is 0.60 Å (similar to that for Li^+, 0.59 Å), whereas that for Cd^{2+} is 0.78 Å (halfway to that for Na^+, 0.99 Å).[10]

To determine the structures formed *via* the interactions of zinc and cadmium with arginine, IRMPD spectroscopy in conjunction with theoretical calculations is utilized in the present study. IRMPD spectroscopy has been employed by various research groups to evaluate the structural characteristics of different metallated amino acids including alkali metal cation complexes with glutamine,[11,12] tryptophan,[13] lysine,[14] arginine,[8] aspartic and glutamic acid,[15] serine,[16] threonine,[17] cysteine,[18] methionine,[19] and histidine,[20] as well as zinc and cadmium complexes of phenylalanine,[21] histidine,[22] proline,[23] cysteine,[24] glutamine,[25] serine,[26] asparagine,[27] glutamic acid,[28] aspartic acid,[29] and lysine.[30] As in these previous studies, the present study determines structures of these complexes by comparing measured IRMPD action spectra of the zinc and cadmium complexes of Arg formed experimentally to calculated IR spectra of predicted low-energy conformers of these complexes. The present study will further the collective knowledge of how biologically important metals interact with amino acid systems, providing fundamental information useful in understanding the nature of metal binding in larger more biologically relevant systems.

Experimental and theoretical methods

Mass spectrometry and photodissociation

IRMPD experiments were conducted at the Free Electron Lasers for Infrared Experiments (FELIX) facility at Radboud University in The Netherlands.[31] Described elsewhere, a 4.7 T Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer was used to measure the IRMPD action spectra.[21,32,33] An electrospray ionization (ESI) source was used to generate ions that were then accumulated in a hexapole trap for 2–3 seconds. The ions were then pulse extracted through a quadrupole deflector and sent through a radiofrequency (RF) octopole ion guide into the ICR cell. Switching the dc bias of the octopole was used to avoid collisional heating of the ions and decrease their potential energy for more efficient trapping within the ICR cell.[21] Once stored within the ICR cell, the ion of interest was assumed to be roughly at room temperature and was mass selected using a stored waveform inverse Fourier transform (SWIFT) excitation pulse.[34,35] The trapped ions were then irradiated with the FELIX laser for 0.5–1 seconds at a 10 Hz macropulse repetition rate with an energy of 20–55 mJ per pulse and a bandwidth of 0.5% of the central frequency. IRMPD action spectra were created by plotting the photofragmentation yield as a function of the frequency of the IR radiation. The photofragmentation yield is defined as $Y = \sum I_F/(\sum I_P + \sum I_F)$, where I_P and I_F are the integrated intensities of the precursor and fragment ion mass peaks (the sum includes all isotopes), respectively. The yield was linearly corrected for the frequency dependent variation in the laser pulse energy. The IR frequency was calibrated online using a grating spectrometer.

The metallated arginine complexes were prepared from 1 mM Arg and 1 mM $Zn(NO_3)_2$ or $CdCl_2$ in 60 : 40 MeOH/H_2O using a Micromass Z-Spray ESI source. This resulted in the formation of $[Zn(Arg–H)]^+$ and $CdCl^+(Arg)$ species, respectively. Notably, species analogous to those observed previously in our Zn^{2+}/Cd^{2+} His study were not observed, including $Zn^{2+}(Arg)_2$ and $Cd^{2+}(Arg)_2$, nor $[Cd(Arg–H)]^+$ nor $M^+(Arg–H)(Arg)$. (The failure to form these species is unfortunate, as they would have allowed for a better comparison to our previous His study and between the metals; however, these results also point to a stronger interaction of Zn and Cd with His than with Arg in solution.) A flow rate of 5 μL min^{-1} was used for both metallated complexes and the ESI needle was held at a voltage of 3 kV. Both the arginine (≥99%) and zinc nitrate hexahydrate (≥99%) samples were purchased from Sigma Aldrich and the cadmium chloride (99.9%) sample was purchased from J. T. Baker Chemical Co. All samples were used as received.

Computational details

To determine the lowest lying conformers of the metallated Arg complexes, we started with structures similar to those from previous results of alkali metal cations binding to arginine from Williams and co-workers, who used molecular dynamics to thoroughly search conformational space (360 total structures).[8,36] For the zinc complexes, we deprotonated all reasonable binding sites (backbone amine, carboxylic acid, and nitrogens on the guanidine group) within Arg. The backbone dihedral angles were then oriented so that the binding sites would be able to efficiently bond to the metal cation and reduce steric hindrance. This process was repeated so that every conceivable bonding orientation and motif was examined. A similar procedure was used for the $CdCl^+$ complexes of neutral Arg. Initial geometry optimizations of the Zn structures were conducted at the B3LYP/6-31G(d,p) level of theory and at the B3LYP/def2-TZVP level for Cd species using the Gaussian09 suite of programs.[37] Def2-TZVP is a balanced basis set for all atoms and includes a triple zeta valence basis set plus polarization functions and a small-core 28 electron effective core potential (ECP) for Cd.[38] The ECP, def2-TZVP, and def2-TZVPP basis sets were acquired from the EMSL basis set exchange.[39] The "loose" keyword was used to promote convergence by implementing a step size of 0.01 a.u. and an rms force constant of 0.0017. Further optimization of the lowest lying metallated

structures was done at the B3LYP/6-311+G(d,p) level for Zn complexes. Geometry optimizations of all low energy structures were also done including empirical dispersion corrections at the B3LYP-GD3BJ level, using the same basis sets noted above.[40]

Vibrational frequencies were calculated at the B3LYP/6-311+G(d,p) level of theory for the Zn species and the B3LYP/def2-TZVP level of theory for the Cd species, respectively, and were scaled by 0.975 to compare them to the IRMPD action spectra. This scaling factor has been proven in the past to account well for any known inaccuracies in the calculated frequencies and thus gives a good comparison between well resolved theoretical and experimental peaks.[24–28] Additionally, the calculated frequencies were broadened using a 20 cm^{-1} full-width at half maximum Gaussian line shape for comparison with the experimental IRMPD data. This broadening accounts for the finite laser bandwidth, unresolved rotational structure of the ions (which are near room temperature), vibrational mode anharmonicity, and broadening as a direct result from the multiple photon absorption process.[41]

Relative energies of the B3LYP geometries located were calculated at the B3LYP, B3P86, and MP2(full) levels using the 6-311+G(2d,2p) and def2-TZVPP basis sets for Zn and Cd complexes, respectively. These same basis sets were also used for single point energies at the B3LYP-GD3BJ level using B3LYP-GD3BJ geometries. 0 K relative enthalpies were calculated by applying zero-point energy (ZPE) corrections to single point energies. Thermal corrections to 298 K Gibbs energies were calculated by using the rigid rotor/harmonic oscillator approximation. ZPE and thermal corrections were calculated by scaling vibrational frequencies by 0.989.[42]

Results and discussion

Theoretical results: nomenclature

Complexes are designated by their metal binding site in brackets, where a negative sign indicates a site of deprotonation (if applicable). Backbone binding sites are designated as follows: N, backbone amine; CO, carbonyl oxygen; CO_2^-, both carboxylate oxygens. Side-chain nitrogens adopt the following nomenclature: the nitrogen closest to the backbone is designated as N^{δ}, the doubly protonated nitrogen on the side chain is designated N^{ω}, and the side-chain nitrogen doubly bound to the carbon atom is represented as $N^{\omega\prime}$. In order to distinguish between similar binding motifs, the dihedral angles describing the amino acid orientation are listed directly after the bracket. These orientations use designations of *cis* (*c*, angles between 0° and 45°), *gauche* (*g*, angles between 45° and 135°), and *trans* (*t*, angles between 135° and 180°). Dihedral angles are measured starting from the carboxylic acid hydrogen (unless this site is deprotonated) or the analogous proton on NH_2 in zwitterionic structures and go to the terminal side-chain $N^{\omega\prime}$ nitrogen (which is usually bound to zinc), followed by the $\angle N^{\delta}CN^{\omega\prime}H$ dihedral angle. For structures with otherwise identical designations, the signs of the *gauche* angles (+ or −) are used to differentiate between them.

Theoretical results: $[Zn(Arg–H)]^+$ relative energies

Low-energy conformers of $[Zn(Arg–H)]^+$ are shown in Fig. 1. Table 1 gives the 0 and 298 K Gibbs energies of these conformers and several others calculated at the B3LYP, B3LYP-GD3BJ, B3P86, and MP2(full) levels of theory relative to the lowest-energy conformer, the ground structure (GS), at each level of theory. The $[Zn(Arg–H)]^+$ GS adopts a tridentate structure, characterized by a $[N,CO^-,N^{\omega\prime}]$ binding motif, where the metal center is bound to the backbone amine nitrogen, the deprotonated carbonyl oxygen, and the terminal side-chain $N^{\omega\prime}$ nitrogen. We located four conformers that take this form, with *gggtgct* as the GS, *ggggggt* and *gcggttt* structures lying 22–25 kJ mol^{-1} higher in energy, and *gggtgcc* about 50 kJ mol^{-1} higher in energy, Table 1. Close examination of these structures indicates that the Zn–guanidine group in the GS is more planar (*e.g.*, $\angle N^{\omega}CN^{\omega\prime}Zn = 155°$) compared to the higher energy conformers (146°, 150°, and 93°, respectively). (The *gcggttt* structure is higher in energy than *ggggggt* because it has the $N^{\delta}H$ bond pointed towards Zn, rather than away, as in the three other structures.) This planarity causes a reduction in steric hindrance, leading to more energetically favorable structures. The *gggtgcc* structure lies highest in energy because it has the hydrogen on the $N^{\omega\prime}$ binding site in a *cis* orientation, compared to the *trans* orientation in the otherwise comparable *gggtgct* GS. This orients this hydrogen more towards the Zn ($\angle ZnN^{\omega\prime}H = 107°$) compared to $\angle ZnN^{\omega\prime}H = 119°$ for the GS.

Three other tridentate binding motifs were discovered, $[CO_2^-,N^{\omega\prime}]$, $[N^-,CO,N^{\omega\prime}]$, and $[N^-,OH,N^{\omega\prime}]$. The lowest energy $[CO_2^-,N^{\omega\prime}]$ structure is *gggtgct* (39–48 kJ mol^{-1} above the GS), Fig. 1. Four other orientations include *gcgtgct*, *ggggttt*, and two forms of *gggtgcc* (differing only in the backbone NH_2 orientation) lying 44–53, 44–55, 75–86, and 83–92 kJ mol^{-1} higher in energy than the GS, respectively. Generally lying higher in energy than both of these binding motifs are the $[N^-,CO,N^{\omega\prime}]$ structures (*tgggtgct*, *tgcggttt*, *tgggtgcc*, and *tggttcct*), which are 39–44, 61–65, 96–101, and 127–133 kJ mol^{-1} higher in energy than the GS, respectively. The $[N^-,OH,N^{\omega\prime}]$ structures (*tgggtgct* and *tgcggttt*), in which the zinc binds to the hydroxyl oxygen rather than the carbonyl oxygen, lie higher still, approximately 36–52 kJ mol^{-1} higher than the analogous $[N^-,CO,N^{\omega\prime}]$ structures. The same is true for *tgggtgcc* and *tggttcct* analogues, not included in Table 1 (see ESI,† Table S1). For both of these binding motifs, deprotonation of the amine instead of the more acidic carboxylic acid accounts for their relatively high energies.

Tetradentate and bidentate binding motifs were also found, but are unlikely to be populated experimentally because of their high energies (>107 kJ mol^{-1} above the GS, see Table 1 and Table S1, ESI†). Select higher energy species are shown in the ESI,† Fig. S1. It is interesting to note that the lowest energy orientation of all three tridentate binding motifs is *gggtgct*. This orientation appears to be the most efficient at reducing ring strain, reducing steric effects, and allowing planarity of the guanidine group. However, the *gggtgct* orientation is not found in the tetradentate or bidentate binding motifs and might contribute to their higher energies.

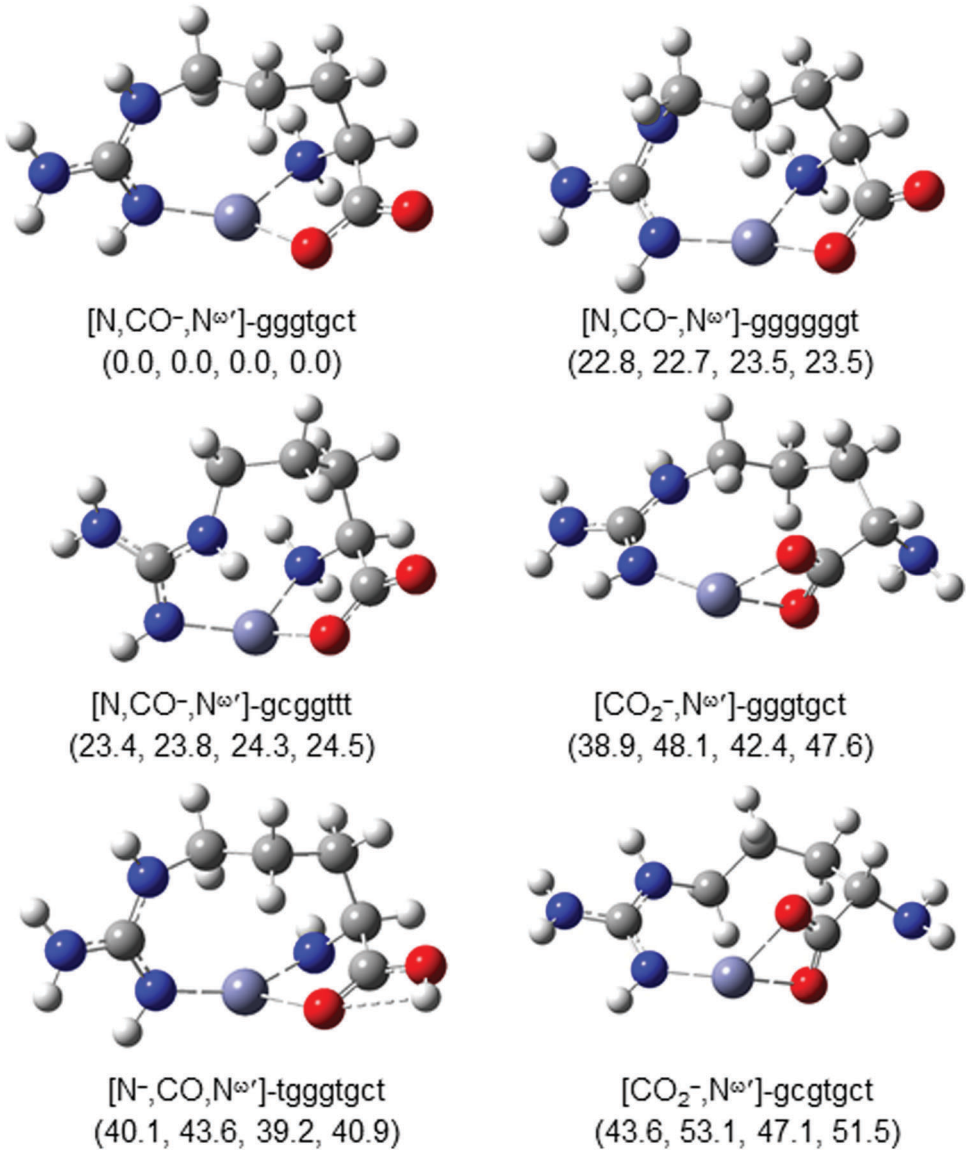

Fig. 1 Structures of low energy $[Zn(Arg-H)]^+$ conformers calculated at the B3LYP/6-311+G(d,p) level of theory. Relative 0 K enthalpies (kJ mol^{-1}) are given at the B3LYP, B3LYP-GD3BJ, B3P86, and MP2(full) levels, respectively. Short dashed lines indicate hydrogen bonds and long dashed lines indicate Zn–ligand bonds.

Theoretical results: $[Zn(Arg-H)]^+$ structures

Full descriptions of key geometric parameters of low-energy $[Zn(Arg-H)]^+$ conformers are given in Table 2. The Zn–O bond distance for the $[N,CO^-,N^{\omega\prime}]$-*gggtgct* GS is 1.88 Å, and those for the *ggggggt*, *gcggttt*, and *gggtgcc* conformers are within 0.01 Å. Additionally, the Zn–N bond lengths of all four of these conformers are within 0.02 Å of one another (~2.09 Å), as are the Zn–$N^{\omega\prime}$ bond lengths, which are shorter by about 0.17 Å (~1.92 Å). Zn–O bond distances within the $[N,CO^-,N^{\omega\prime}]$ structures are shorter (by an average of 0.35 Å) than those seen in the $[N^-,CO,N^{\omega\prime}]$ complexes. Likewise, deprotonation at the amino nitrogen site affects the Zn–N bond distances observed, where Zn–N bond distances in the $[N^-,CO,N^{\omega\prime}]$ and $[N^-,OH,N^{\omega\prime}]$ binding motifs are comparable to the Zn–O bond distances in the $[N,CO^-,N^{\omega\prime}]$ structures. Notably, for the $[CO_2^-,N^{\omega\prime}]$ structures, the Zn–O bond distances increase to a range of 1.98–2.08 Å because the two oxygens share the delocalized charge. Indeed, these two bond distances are nearly identical in all but the *ggggttt* conformer, where one of the oxygens forms a hydrogen bond with $N^{\delta}H$, lengthening its bond to zinc. Calculated Zn–$N^{\omega\prime}$ bond distances are similar for all located conformers (within 0.06 Å) and more similar throughout each binding motif, with differences less than 0.02 Å. This observation is explained by the fact that the guanidinium group almost always complexes at the neutral $N^{\omega\prime}$ lone pair, which leads to zinc binding characterized by similar bond lengths regardless of the other two binding sites. The flexibility of the long Arg side chain facilitates optimal binding at this position.

Table 1 Relative enthalpies (0 K) and Gibbs energies (298 K) of $[Zn(Arg-H)]^+$ species[a]

Structure	B3LYP	B3LYP-GD3BJ[b]	B3P86	MP2(full)
$[N,CO^-,N^{\omega\prime}]$-*gggtgct*	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
$[N,CO^-,N^{\omega\prime}]$-*ggggggt*	22.8 (23.5)	22.7 (23.8)	23.5 (24.3)	23.5 (24.3)
$[N,CO^-,N^{\omega\prime}]$-*gcggttt*	23.4 (22.6)	23.8 (23.5)	24.3 (23.6)	24.5 (23.8)
$[CO_2^-,N^{\omega\prime}]$-*gggtgct*	38.9 (36.9)	48.1 (46.1)	42.4 (40.3)	47.6 (45.6)
$[N^-,CO,N^{\omega\prime}]$-*tgggtgct*	40.1 (39.5)	43.6 (43.3)	39.2 (38.6)	40.9 (40.3)
$[CO_2^-,N^{\omega\prime}]$-*gcgtgct*	43.6 (41.7)	53.1 (51.3)	47.1 (45.3)	51.5 (49.6)
$[CO_2^-,N^{\omega\prime}]$-*ggggttt*	44.1 (42.9)	54.6 (53.4)	48.3 (47.0)	52.5 (51.2)
$[N,CO^-,N^{\omega\prime}]$-*gggtgcc*	49.6 (50.1)	47.1 (48.0)	50.7 (51.2)	50.9 (51.5)
$[N^-,CO,N^{\omega\prime}]$-*tgcggttt*	61.4 (60.9)	64.7 (64.5)	63.3 (62.8)	64.7 (64.2)
$[CO_2^-,N^{\omega\prime}]$-*gggtgcc*[c]	75.1 (73.8)	83.0 (81.6)	80.3 (79.0)	86.2 (84.9)
$[N^-,OH,N^{\omega\prime}]$-*tgggtgct*	80.2 (77.7)	75.7 (73.6)	83.0 (80.5)	77.7 (75.2)
$[CO_2^-,N^{\omega\prime}]$-*gggtgcc*[c]	83.4 (81.5)	91.2 (89.0)	88.1 (86.2)	92.4 (90.5)
$[N^-,CO,N^{\omega\prime}]$-*tgggtgcc*	95.8 (95.1)	96.8 (95.6)	97.3 (96.7)	101.3 (100.7)
$[N^-,OH,N^{\omega\prime}]$-*tgcggttt*	104.8 (100.9)	100.8 (97.8)	110.6 (106.7)	105.1 (101.2)
$[N,CO,N^{\delta-},N^{\omega}]$-*tggggtct*	111.8 (110.1)	114.7 (112.0)	111.1 (109.4)	107.1 (105.4)
$[N^-,CO,N^{\omega\prime}]$-*tggttcct*	130.2 (129.3)	133.4 (132.7)	126.7 (125.9)	129.5 (128.6)

[a] Single point enthalpies and Gibbs energies (in parentheses) calculated at the B3LYP, B3P86, and MP2(full)/6-311+G(2d,2p) levels of theory using B3LYP/6-311+G(d,p) geometries and zero point energy corrections. [b] B3LYP-GD3BJ/6-311+G(2d,2p) energies using B3LYP-GD3BJ/6-311+G(d,p) geometries and zero point energy corrections. [c] These two structures are distinguished by the orientation of the NH_2 group, which hydrogen bonds to a carboxylate oxygen in the lower energy structure (see ESI, Fig. S1).

Table 2 Bond distances (Å) and bond angles (deg) of $[Zn(Arg-H)]^+$ species[a]

Structure	r(Zn–X)	r(Zn–O)	r(Zn–Y_s)	∠XZnO	∠XZnY_s	∠OZnY_s
$[N,CO^-,N^{\omega\prime}]$-*gggtgct*	2.10	1.88	1.92	85.5	129.5	140.6
$[N,CO^-,N^{\omega\prime}]$-*ggggggt*	2.09	1.87	1.92	86.2	116.7	151.7
$[N,CO^-,N^{\omega\prime}]$-*gcggttt*	2.08	1.89	1.94	87.3	120.5	133.5
$[CO_2^-,N^{\omega\prime}]$-*gggtgct*	2.01[b]	2.01	1.91	66.0[b]	129.9[b]	135.5
$[N^-,CO,N^{\omega\prime}]$-*tgggtgct*	1.85	2.19	1.92	82.3	151.1	115.7
$[CO_2^-,N^{\omega\prime}]$-*gcgtgct*	2.01[b]	2.01	1.91	66.1[b]	135.7[b]	129.2
$[CO_2^-,N^{\omega\prime}]$-*ggggttt*	1.98[b]	2.08	1.91	65.3[b]	131.6[b]	115.0
$[N,CO^-,N^{\omega\prime}]$-*gggtgcc*	2.10	1.87	1.93	86.2	116.2	141.9
$[N^-,CO,N^{\omega\prime}]$-*tgcggttt*	1.86	2.21	1.94	83.6	135.0	111.7
$[CO_2^-,N^{\omega\prime}]$-*gggtgcc*	2.04[b]	2.01	1.93	65.7[b]	118.7[b]	120.7
$[N^-,OH,N^{\omega\prime}]$-*tgggtgct*	1.83	2.33[c]	1.90	77.4[c]	159.2	113.0[c]
$[CO_2^-,N^{\omega\prime}]$-*gggtgcc*	2.02[b]	2.02	1.93	65.8[b]	121.9[b]	118.4
$[N^-,CO,N^{\omega\prime}]$-*tgggtgcc*	1.88	2.16	1.95	85.8	127.9	117.3
$[N^-,OH,N^{\omega\prime}]$-*tgcggttt*	1.83	2.51[c]	1.91	77.2[c]	146.2	105.5[c]
$[N,CO,N^{\delta-},N^{\omega}]$-*tggggtct*	2.11	2.11	1.92[d]	79.3	103.7[d]	108.7[d]
$[N^-,CO,N^{\omega\prime}]$-*tggttcct*	1.90	2.34	1.96	81.5	150.0	112.6

[a] Unless otherwise noted, X = backbone amine, O = backbone carbonyl, Y_s = side-chain nitrogen ($N^{\omega\prime}$). [b] X = second oxygen on CO_2^- group. [c] O = hydroxyl oxygen on backbone. [d] Y_s = side-chain nitrogen (N^{δ}). In this structure, r(Zn–N^{ω}) = 2.09 Å, ∠N^{δ}ZnN^{ω} = 155.4°, and ∠OZnN^{ω} = 125.2°.

Table 2 also shows that the bond angle associated with ∠XZnO (where X = amino nitrogen in CS structures and the second carboxylate oxygen in the SB structures) metal binding varies with binding motif, but does not change significantly within a specific bonding motif, *i.e.*, upon conformational changes in the dihedral angles along the backbone. For example, within the $[N,CO^-,N^{\omega\prime}]$ species, the ∠NZnO angle is 86.4 ± 0.9°. A similar range is observed for the zwitterionic $[CO_2^-,N^{\omega\prime}]$ species, where ∠OZnO bonding angles are calculated as 65.7 ± 0.4°. (These more shallow ∠OZnO angles are a result of the four-membered CO_2^-–Zn ring in $[CO_2^-,N^{\omega\prime}]$ structures rather than a five-membered C–CO^-–Zn–N ring observed in $[N,CO^-,N^{\omega\prime}]$.) If the amino group (rather than the carboxylic acid) is deprotonated, then the ∠NZnO angle of the four $[N^-,CO,N^{\omega\prime}]$ conformers drops slightly to 83.7 ± 2.2°, and yet again to ~77° for $[N^-,OH,N^{\omega\prime}]$ conformers. These similarities are a result of the rigidity of the arginine backbone. In contrast, bond angles involving the side chain vary much more broadly (104°–159°), indicative of the different dihedral angles associated with the different conformations.

Theoretical results: $CdCl^+$(Arg) relative energies

As shown in Table 3, theory does not agree on the GS for the $CdCl^+$(Arg) complex. One possibility is a bidentate $[N,CO^-](N^{\omega\prime}H_2^+)$-*gggtgc* conformer, where the carboxylic acid hydrogen has migrated to the side-chain nitrogen, $N^{\omega\prime}$, as represented by $N^{\omega\prime}H_2^+$ in parentheses, forming a salt-bridge species. (Note the ∠$N^{\delta}CN^{\omega\prime}H$ dihedral angle designation is excluded as there are two hydrogens on $N^{\omega\prime}$.) Another possibility is a tridentate $[N,CO,N^{\omega\prime}]$-*tgggtgct* structure. These structures and other low energy conformers can be seen in Fig. 2. The $[N,CO^-](N^{\omega\prime}H_2^+)$ structure is lower in energy than $[N,CO,N^{\omega\prime}]$-*tgggtgct* by 9.4 and 5.5 kJ mol^{-1} at the B3LYP and B3P86 levels of theory; however, at the B3LYP-GD3BJ and MP2(full) levels of theory, the

Table 3 Relative enthalpies (0 K) and Gibbs energies (298 K) of $CdCl^+$(Arg) species[a]

Structure	B3LYP	B3LYP-GD3BJ[b]	B3P86	MP2(full)
[N,CO$^-$](N$^{\omega\prime}$H$_2^+$)-*gggtgc*	0.0 (0.0)	3.0 (2.2)	0.0 (0.0)	12.6 (11.1)
[N,CO,N$^{\omega\prime}$]-*tgggtgct*	9.4 (10.8)	0.0 (0.0)	5.5 (7.0)	0.0 (0.0)
[N,CO,N$^{\omega\prime}$]-*tgtgggct*	21.6 (23.4)	12.2 (11.8)	17.8 (19.6)	12.2 (12.6)
[N,CO,N$^{\omega\prime}$]-*tgggggct*	33.5 (34.2)	24.7 (24.3)	30.7 (31.3)	23.8 (22.9)
[N,OH,N$^{\omega\prime}$]-*tgggtgct*	40.3 (39.6)	28.6 (24.1)	38.3 (37.7)	29.3 (27.2)
[N,OH,N$^{\omega\prime}$]-*tgtgggct*	45.7 (45.6)	27.1 (26.4)	44.0 (43.9)	34.9 (33.3)
[N,CO,N$^{\omega\prime}$]-*tgcgggcc*	49.6 (51.9)	37.0 (39.3)	47.2 (49.6)	39.4 (40.2)
[CO,N$^{\omega\prime}$]-*ctgttgcc*	49.2 (46.9)	54.2 (49.6)	48.6 (46.3)	58.8 (55.1)
[N,N$^{\omega\prime}$]-*ttgtggct*	55.5 (54.2)	37.4 (36.6)	54.1 (52.7)	49.9 (47.1)
[CO,N$^{\omega\prime}$]-*tg$_-$g$_-$g$_+$tg$_+$ct*[c]	56.9 (58.9)	51.3 (52.6)	55.9 (57.9)	55.4 (55.9)
[CO,N$^{\omega\prime}$]-*tg$_+$g$_+$g$_-$tg$_-$ct*	60.6 (60.5)	56.7 (56.1)	60.9 (60.8)	63.9 (62.3)
[CO,N$^{\omega\prime}$]-*ttgttgct*	60.9 (57.4)	58.2 (54.9)	62.2 (58.7)	63.2 (58.4)
[CO,N$^{\omega\prime}$]-*ttgttgcc*[cd]	63.2 (63.5)	61.3 (62.7)	64.4 (64.7)	68.4 (67.3)
[CO,N$^{\omega\prime}$]-*ttgttgcc*[d]	65.6 (62.9)	66.1 (64.7)	68.5 (65.9)	71.9 (67.8)
[N$^{\omega\prime}$]-*tgggtgtc*	66.2 (62.0)	63.1 (62.0)	67.9 (63.8)	74.5 (68.9)
[N,CO,N$^{\omega\prime}$]-*tggttcct*	70.8 (71.4)	62.5 (61.7)	65.2 (65.8)	59.8 (58.9)
[N,OH,N$^{\omega\prime}$]-*tgcgggcc*	71.4 (75.1)	44.8 (49.9)	68.8 (72.6)	53.3 (55.6)
[N$^{\omega\prime}$]-*ttgttgtc*	94.4 (85.7)	86.0 (81.5)	100.4 (91.7)	113.5 (103.4)
[N$^{\omega\prime}$]-*ttggtgcc*	99.9 (91.0)	72.4 (67.5)	106.3 (97.4)	118.4 (108.0)
[OH,N$^{\omega\prime}$]-*tcgttgct*	105.3 (101.1)	93.0 (89.5)	108.9 (104.7)	110.3 (104.6)
[N$^{\omega\prime}$]-*tcggtgcc*	114.5 (104.9)	133.1 (121.5)	121.3 (117.1)	133.8 (122.7)

[a] Single point enthalpies and Gibbs energies (in parentheses) calculated at the B3LYP, B3P86, and MP2(full)/def2-TZVPP levels of theory using B3LYP/def2-TZVP geometries and zero point energy corrections. [b] B3LYP-GD3BJ/def2-TZVPP energies using B3LYP-GD3BJ/def2-TZVP geometries and zero point energy corrections. [c] This structure has a OH•Cl hydrogen bond. [d] These structures are distinguished by the orientation of the backbone NH_2 group.

[N,CO$^-$](N$^{\omega\prime}$H$_2^+$) complex is 3.0 and 12.6 kJ mol^{-1} higher in energy than [N,CO,N$^{\omega\prime}$]-*tgggtgct*, respectively. Four other [N,CO,N$^{\omega\prime}$] structures with orientations of *tgtgggct*, *tgggggct*, *tgcgggcc*, and *tggttcct* were found and lie ~12, 24–29, 37–42, and 60–63 kJ mol^{-1} higher in energy than the [N,CO,N$^{\omega\prime}$]-*tgggtgct* conformer.

Tridentate structures are not as common as those found in $[Zn(Arg–H)]^+$ species with the [N,OH,N$^{\omega\prime}$] motif (*tgggtgct*, *tgtgggct*, and *tgcgggcc*) being the only other tridentate structure found. (The [N,OH,N$^{\omega\prime}$] analogue of [N,CO,N$^{\omega\prime}$]-*tgggggct* collapsed to a [N,N$^{\omega\prime}$]-*ttgtggct* structure, where there is no longer a Cd–OH interaction.) Compared to their analogous [N,CO,N$^{\omega\prime}$] complexes, these structures lie 28–33, 15–26, and 8–22 kJ mol^{-1} higher in energy, a result of the lower cadmium cation binding affinity for the hydroxyl group on the carboxylic acid compared to the carbonyl.

In addition to the [N,CO$^-$](N$^{\omega\prime}$H$_2^+$)-*gggtgc* structure, bidentate structures include the [N,N$^{\omega\prime}$]-*ttgtggct* structure noted above, as well as six versions of [CO,N$^{\omega\prime}$], and [OH,N$^{\omega\prime}$]-*tcgttgct*. Select higher energy species, including bidentate and monodentate structures, are shown in the ESI,† Fig. S2. The six bidentate [CO,N$^{\omega\prime}$] complexes were found to be 49–72 kJ mol^{-1} higher in energy than the GS with the *ctgttgcc* orientation being the lowest in energy and *ttgttgcc* being the highest. No tetradentate structures were found suggesting that this binding conformation for neutral Arg is unfavorable energetically. Four monodentate species were found with [N$^{\omega\prime}$]-*tgggtgtc*, *ttgttgtc*, *ttggtgcc*, and *tcggtgcc* being 63–75, 86–114, 72–118, and 114–134 kJ mol^{-1} higher in energy than the GS, respectively.

Theoretical results: $CdCl^+$(Arg) structures

Table 4 shows the binding geometries, bond distances, and dihedral angles of each of the $CdCl^+$(Arg) complexes found. The [N,CO$^-$](N$^{\omega\prime}$H$_2^+$)-*gggtgc* complex is unique from the other bidentate structures because of its zwitterionic character. In this structure, the metal cation binds at the backbone amino group and one oxygen of the deprotonated carboxylate group, and is stabilized by a hydrogen bond between the uncoordinated carbonyl oxygen and the protonated N$^{\omega\prime}$, Fig. 2. The partial anionic charge on the carboxylate leads to the shortest Cd–O bond distance, 0.23 Å shorter than the next shortest one. The Cd–N bond distance of the [N,CO$^-$](N$^{\omega\prime}$H$_2^+$) complex is slightly shorter (by 0.01–0.13 Å) than those calculated for the [N,CO,N$^{\omega\prime}$] binding motifs.

The Cd–N bond distance of the [N,CO,N$^{\omega\prime}$]-*tgggtgct* conformer is 2.44 Å, and those of the *tgtgggct*, *tgggggct*, *tgcgggcc*, and *tggttcct* orientations are within 0.07 Å. Cd–O and Cd–N$^{\omega\prime}$ bond distances of all five of these conformers are within 0.11 and 0.08 Å of one another, respectively. The angles calculated in these structures vary extensively, with the exception of ∠NCdO, which is 68.5 ± 2.3°. The ∠NCdN$^{\omega\prime}$ angle varies up to 28° between the different conformers, while ∠OCdN$^{\omega\prime}$ angles differ even more, by up to 36°. Both the [N,CO,N$^{\omega\prime}$] and [N,OH,N$^{\omega\prime}$] structures have similar Cd–N (averages of 2.44 and 2.40 Å, respectively) and Cd–N$^{\omega\prime}$ (2.22 and 2.17 Å, respectively) bond distances because the local binding environments have not changed. Cd–O distances in [N,OH,N$^{\omega\prime}$] conformers are on average 0.16 Å longer than in [N,CO,N$^{\omega\prime}$] structures. The ∠NCdO and ∠OCdN$^{\omega\prime}$ dihedral angles for both binding motifs are very similar with variations up to 6.5° and 7.9°, respectively. The ∠NCdN$^{\omega\prime}$ angle of [N,OH,N$^{\omega\prime}$] conformers is on average 13.3° larger than those calculated in the [N,CO,N$^{\omega\prime}$] structures.

Another major binding motif is bidentate [CO,N$^{\omega\prime}$], with the *ctgttgcc* orientation being the lowest in energy. These other bidentate structures share few structural similarities with the

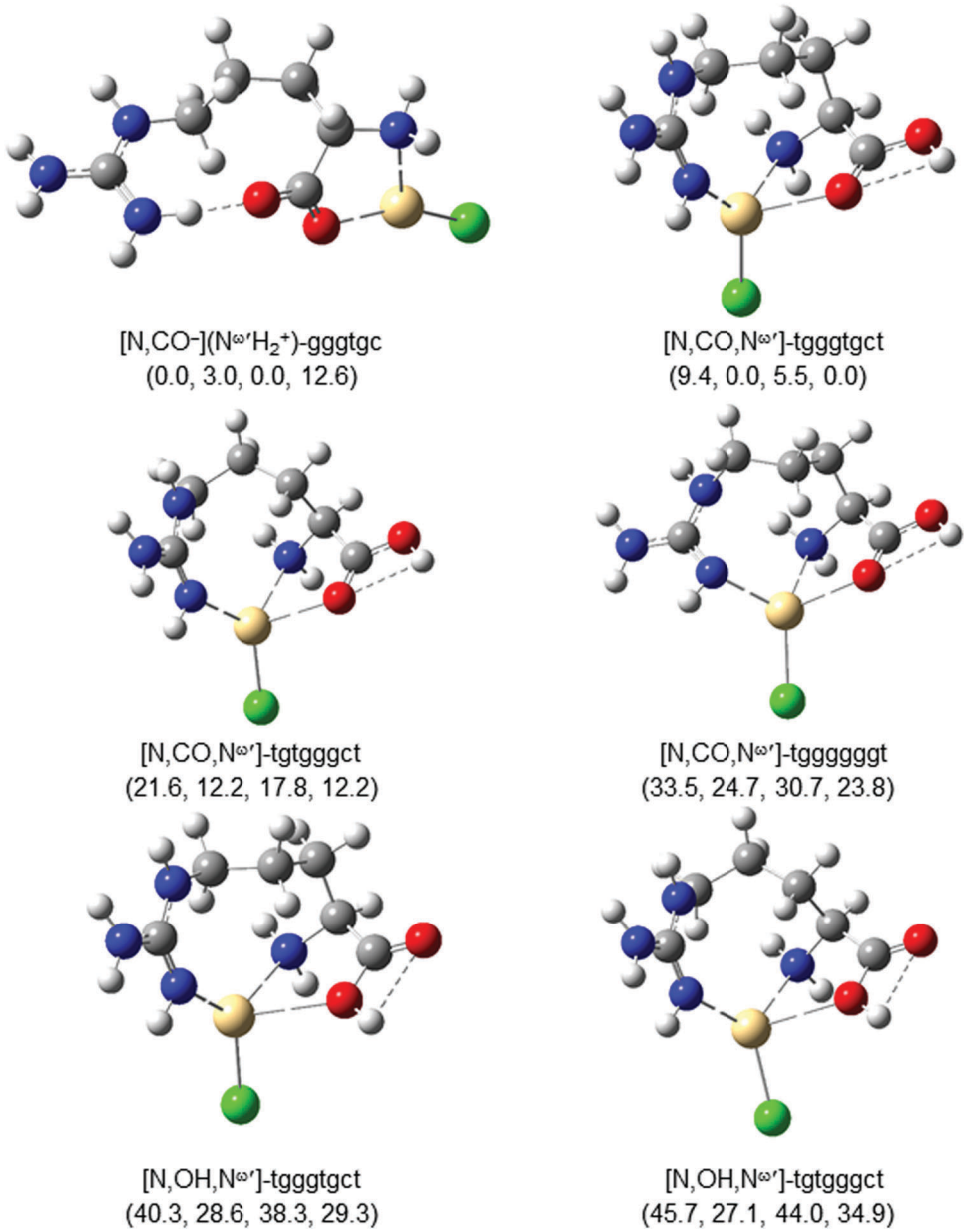

Fig. 2 Structures of low energy $CdCl^+$(Arg) conformers calculated at the B3LYP/def2-TZVP level of theory. Relative 0 K enthalpies (kJ mol^{-1}) are given at the B3LYP, B3LYP-GD3BJ, B3P86, and MP2(full) levels, respectively. Short dashed lines indicate hydrogen bonds and long dashed lines indicate Cd–ligand bonds.

bidentate [N,CO^-]($N^{\omega\prime}H_2^+$) GS structure. Because of the difference in charge state on the oxygen, the Cd–O bond length of the lowest energy [CO,$N^{\omega\prime}$] structure is 0.29 Å longer than in the [N,CO^-]($N^{\omega\prime}H_2^+$) structure. Cd–O and Cd–$N^{\omega\prime}$ bond distances in the [CO,$N^{\omega\prime}$] conformers (2.35–2.47 and 2.11–2.13 Å, respectively) exhibit small variances, and are comparable to bond lengths observed in the tridentate species. For the monodentate [$N^{\omega\prime}$] conformers, the Cd–$N^{\omega\prime}$ bond lengths are the shortest observed among all conformers, 2.06–2.09 Å.

IRMPD action spectroscopy

The experimentally measured IRMPD action spectra of $[Zn(Arg–H)]^+$ and $CdCl^+$(Arg) complexes were scanned from 5.4 to 16.7 μm (600–1800 cm^{-1}) and the IR induced fragmentation yield was calculated at each laser frequency using the equation described above. ^{64}Zn is the most abundant isotope (49.17% natural abundance) such that mass to charge ratios (m/z) (given below in brackets) were calculated using this isotope. ^{66}Zn and ^{68}Zn (27.73 and 18.45% natural abundance, respectively)[43] were also monitored and included in the yield calculations for the precursor complex. Zn containing product ions were monitored using ^{64}Zn.

IRMPD of $[Zn(Arg–H)]^+$ precursor complexes resulted in four major products that appeared at m/z 193, 191, 112, and 70, reactions (1)–(4). Representative mass spectra depicting the IRMPD of $[Zn(Arg–H)]^+$ are given in the ESI,† Fig. S3.

Table 4 Bond distances (Å) and bond angles (deg) of $CdCl^+$(Arg) species[a]

Structure	r(Cd–N)	r(Cd–O)	r(Cd–Y_s)	∠NCdO	∠NCdY_s	∠OCdY_s
[N,CO^-]($N^{\omega\prime}H_2^+$)-*gggtgc*	2.38	2.13		76.0		
[N,CO,$N^{\omega\prime}$]-*tgggtgct*	2.44	2.43	2.19	68.9	121.3	107.9
[N,CO,$N^{\omega\prime}$]-*tgtgggct*	2.40	2.47	2.20	68.9	112.7	108.9
[N,CO,$N^{\omega\prime}$]-*tggggggt*	2.39	2.40	2.25	70.8	100.0	117.7
[N,OH,$N^{\omega\prime}$]-*tgggtgct*	2.39	2.62[b]	2.17	64.3[b]	124.2	108.6[b]
[N,OH,$N^{\omega\prime}$]-*tgtgggct*	2.36	2.68[b]	2.18	64.7[b]	115.1	112.6[b]
[N,CO,$N^{\omega\prime}$]-*tgcgggcc*	2.48	2.51	2.18	66.9	97.8	85.3
[CO,$N^{\omega\prime}$]-*ctgttgcc*		2.42	2.13			79.9
[N,$N^{\omega\prime}$]-*ttgtggct*	2.39		2.17		115.8	
[CO,$N^{\omega\prime}$]-*tg*$_-$*g*$_-$*g*$_+$*tg*$_-$*ct*[c]		2.35	2.11			98.2
[CO,$N^{\omega\prime}$]-*tg*$_+$*g*$_+$*g*$_-$*tg*$_+$*ct*		2.36	2.13			101.0
[CO,$N^{\omega\prime}$]-*ttgttgct*		2.42	2.11			83.6
[CO,$N^{\omega\prime}$]-*ttgttgcc*[c,d]		2.40	2.13			98.9
[CO,$N^{\omega\prime}$]-*ttgttgcc*[d]		2.47	2.11			81.5
[$N^{\omega\prime}$]-*tgggtgtc*			2.06			
[N,CO,$N^{\omega\prime}$]-*tggttcct*	2.51	2.49	2.26	66.2	125.5	102.7
[N,OH,$N^{\omega\prime}$]-*tgcgggcc*	2.46	2.71[b]	2.17	59.8[b]	95.8	86.2[b]
[$N^{\omega\prime}$]-*ttgttgtc*			2.07			
[$N^{\omega\prime}$]-*ttggtgcc*			2.07			
[OH,$N^{\omega\prime}$]-*tcgttgct*		2.76[b]	2.09			112.8[b]
[$N^{\omega\prime}$]-*tcggtgcc*			2.08			

[a] Unless otherwise noted, N = amino nitrogen on backbone, O = carbonyl oxygen on backbone, Y_s = side-chain nitrogen ($N^{\omega\prime}$). [b] O = backbone hydroxyl oxygen. [c] This structure has a OH•Cl hydrogen bond. [d] These structures are distinguished by the orientation of the backbone NH_2 group.

$$[Zn(Arg–H)]^+ [237] + nhv \rightarrow Zn(C_5H_{13}N_4)^+ [193] + CO_2 \quad (1)$$

$$\rightarrow Zn(C_5H_{11}N_4)^+ [191] + H_2O + CO \quad (2)$$

$$\rightarrow C_5H_{10}N_3^+ [112] + Zn + CO_2 + NH_3 \quad (3)$$

$$C_5H_{10}N_3^+ [112] + nhv \rightarrow C_4H_8N^+ [70] + HN{=}C{=}NH \quad (4)$$

Once the isotopic distribution is accounted for, the m/z 193 (loss of CO_2), 191 (concomitant loss of H_2O and CO), 112 (loss of the metal center, CO_2, and NH_3), and 70 (loss of CN_2H_2 from m/z 112) fragments were produced with similar intensities. Reaction (4) is assigned as a sequential dissociation process because some of the bands are missing in its appearance spectrum.

IRMPD dissociation of $CdCl^+$(Arg) resulted in only one major fragmentation pathway, the loss of H_2O corresponding to reaction (5). Representative mass spectra depicting the IRMPD of $CdCl^+$(Arg) are given in the ESI,† Fig. S4.

$$CdCl^+(Arg) [323] + nhv \rightarrow CdCl^+(C_6H_{12}N_4O) [305] + H_2O \quad (5)$$

The masses given correspond to the most abundant isotopes of both Cd and Cl, ^{114}Cd (28.73%) and ^{35}Cl (75.76%), such that m/z 323 is designated as the precursor ion. The ^{37}Cl isotope (24.24%) was also monitored and included in yield calculations. Additional ^{113}Cd, ^{112}Cd, ^{111}Cd, and ^{110}Cd (natural abundances of 12.22%, 24.13%, 12.80%, and 12.49%, respectively)[43] isotopes were not monitored.

Overall, these decomposition pathways are similar to those observed previously for metallated Arg complexes. Jockusch *et al.* used thermal radiation and collisional activation to study the effects of alkali metal cation size on the fragmentation pathways of Arg complexes.[36] Notably, they observed that increasing metal cation size changed the structure of the Arg complexes from charge solvated (which mainly lost H_2O) to salt bridges (which mainly lost NH_3). Here, we note that the ionic radius of Cd^{2+}, 0.78 Å, is between that of Li^+ and Na^+ (0.59 and 0.99 Å). Ionic radii of K^+, Rb^+, and Cs^+ are much larger, 1.37, 1.52, and 1.67 Å.[10] This similarity to Li^+ and Na^+ shows that the loss of H_2O in reaction (5) is a reasonable fragment for the $CdCl^+$(Arg) species. Likewise, reaction (2) also includes loss of H_2O for the smaller zinc dication, ionic radius of 0.60 Å, which is similar to that of Li^+. However, for reaction (2), water loss is accompanied by loss of CO, which is comparable to the observations of Lavanant and Hoppilliard for the decomposition of Cu^+(Arg).[44] As for Zn^{2+}, complexation of Cu^{2+} to Arg formed the deprotonated complex, $[Cu(Arg–H)]^+$, which was observed to decompose by loss of both CO_2 and HCO_2,[45] where the former process is analogous to reaction (1). Clearly, the overall energetics of these collision-induced dissociation (CID) experiments differ from those in the IRMPD experiments conducted here, both in terms of the energy deposition as well as the strength of the binding to different metal cations. Loss of CO_2 in reaction (1) is potentially consistent with a deprotonated carboxylic acid, a channel that would not be available in the studies of the M^+(Arg) where M^+ = Li^+–Cs^+, Cu^+, or $CdCl^+$.

Comparison of theoretical and experimental IR spectra: $[Zn(Arg–H)]^+$

The experimental $[Zn(Arg–H)]^+$ spectrum, shown in Fig. 3, has bands located at 632, 1067, 1150, 1246, 1424, 1584, 1639, and 1758 cm^{-1}. It can also be seen that the calculated spectra associated with the lowest energy [N,CO^-,$N^{\omega\prime}$] structures correlate well with the most prominent features of the experimental spectrum. An overview of the predicted vibrational frequencies and intensities of the five lowest energy conformers is given in the ESI,† Table S2. The [N,CO^-,$N^{\omega\prime}$] conformers, especially the

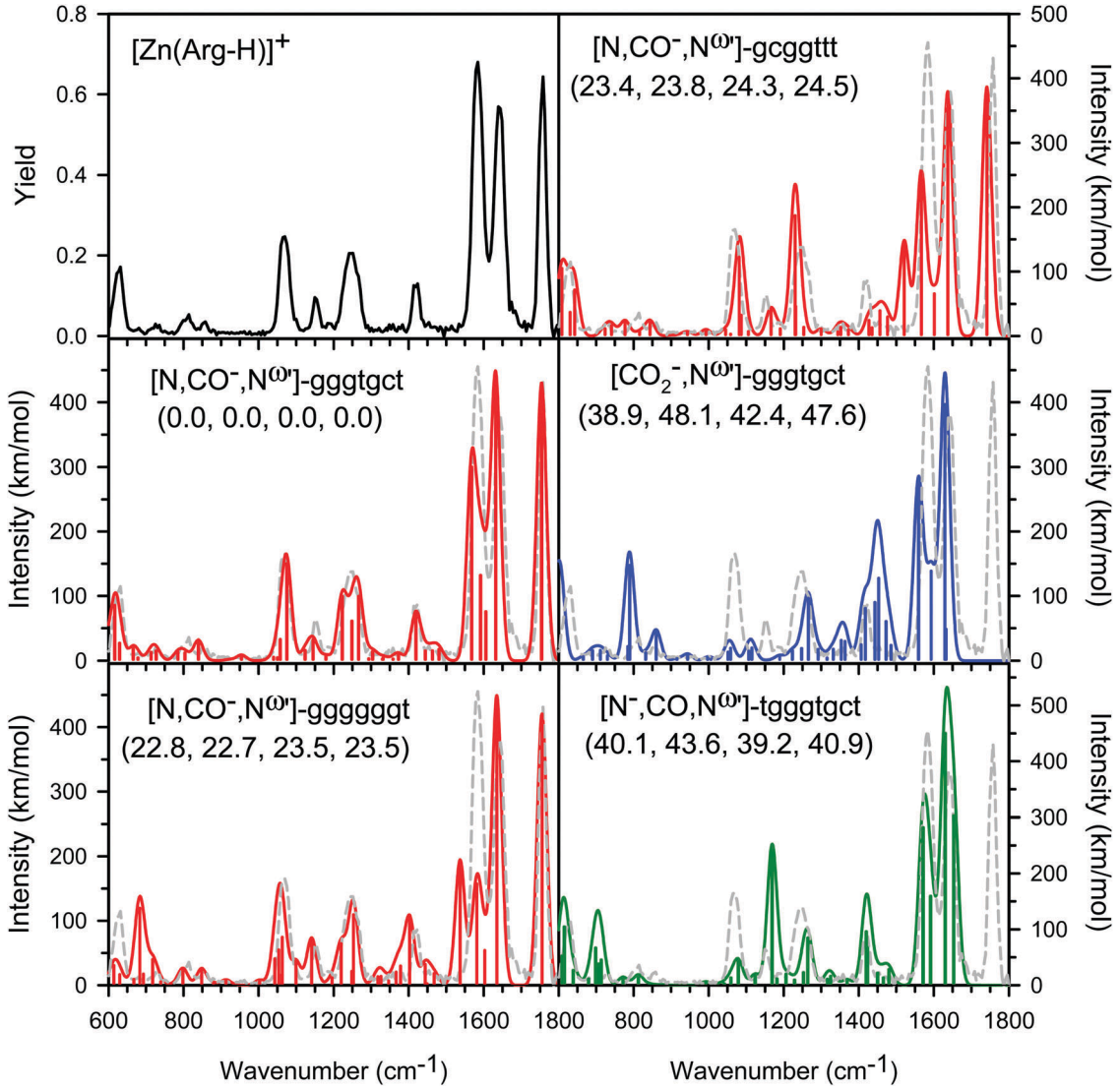

Fig. 3 Comparison of the $[Zn(Arg-H)]^+$ experimental IRMPD action spectrum (black, upper left; dashed grey, otherwise) with IR spectra calculated at the B3LYP/6-311+G(d,p) level of theory for the lowest lying conformers (colored lines). Relative 0 K enthalpies (kJ mol^{-1}) are given at the B3LYP, B3LYP-GD3BJ, B3P86, and MP2(full) levels, respectively.

gggtgct GS, reproduce the highest frequency bands well. Specifically, the *gggtgct* GS predicts bands at 1754 (uncoordinated carbonyl CO stretch), 1632 ($N^{\omega}H_2$ bend), unresolved bands at 1606 and 1591 (backbone NH_2 bend and concerted guanidine NH bends), and 1568 cm^{-1} ($CN^{\omega\prime}$ stretch, $N^{\delta}CN^{\omega}$ bend, and $N^{\omega}H$ bend) that correlate quite well in frequency and intensity to the three highest frequency bands in the experimental spectrum. As shown in Fig. 3, fairly good agreement is also observed for the less intense bands at lower frequency. Here, predicted bands at 1419 ($N^{\omega\prime}H$ and $N^{\delta}H$ bends), unresolved bands at 1266 (mainly $N^{\omega\prime}H$ bend), 1249 (concerted CH_2 twists), and 1223 (concerted CO stretch and CH_2 twists), 1145 (HNCH bend), 1075 (backbone NH_2 wag), and 617 cm^{-1} (out-of-plane $N^{\omega\prime}H$ bend), along with other minor bands near 800 cm^{-1} reproduce the experimental spectrum very well.

The two other low-lying [N,CO^-,$N^{\omega\prime}$] structures, *gggggggt* and *gcggttt*, are consistent with the experimental spectrum throughout most of the region explored, but have significant differences in both cases. The *gggggggt* and *gcggttt* structures predict bands that are in relatively good agreement with the experimental bands located at 1584, 1639, and 1758 cm^{-1} (calculated at 1581, 1635, and 1755 cm^{-1} for *gggggggt* and 1567, 1637, and 1741 cm^{-1} for *gcggttt*); however, they also have intense bands at 1537 and 1521 cm^{-1}, respectively, that do not appear in the experimental spectrum. The 1537 cm^{-1} band in the *gggggggt* spectrum is a combination of CN^{δ} stretch and $N^{\delta}H$, $N^{\omega\prime}H$, and $N^{\omega}H$ bends (which corresponds to the 1567 cm^{-1} band for *gcggttt*). The analogous 1521 cm^{-1} band in the *gcggttt* conformer is a mixture of $N^{\omega}H_2$ and $N^{\omega\prime}H$ bends and CN^{ω} stretch (assigned to the 1581 cm^{-1} band for *gggggggt*). The same two bands in the *gggtgct* GS lie at 1568 and 1591 cm^{-1}, such that they overlap and are not red-shifted compared to experiment. Bands observed between 800 and 1450 cm^{-1} are generally consistent with both excited conformers in shape and intensity,

although the *gcggttt* conformer does not reproduce the band at 1424 cm^{-1} very well. Interestingly, the opposite correlation is observed at the lowest frequencies analyzed. Here, the experimental band at 627 cm^{-1} is better reproduced by the *gcggttt* predicted bands centered at 630 cm^{-1}, whereas the *ggggggt* conformer instead predicts an intense band at 683 cm^{-1}, which is not observed in the IRMPD spectrum.

Fig. 3 also illustrates the predicted IR spectra for alternative binding motifs, [CO_2^-,$N^{\omega\prime}$]-*gggtgct* and [N^-,CO,$N^{\omega\prime}$]-*tgggtgct*. In neither case do the predicted spectra match the experimental features very well. Specifically, both [CO_2^-,$N^{\omega\prime}$] and [N^-,CO,$N^{\omega\prime}$] conformers fail to reproduce the band observed at 1758 cm^{-1} corresponding to the uncoordinated carbonyl stretch. This is clear for [CO_2^-,$N^{\omega\prime}$] as both carboxylate oxygens now coordinate the metal. For [N^-,CO,$N^{\omega\prime}$], the carbonyl is coordinated and the CO that was uncoordinated in the [N,CO^-,$N^{\omega\prime}$] structures is now protonated, *i.e.*, a hydroxyl group. Although these two structures do have predicted bands somewhat consistent with the experimental bands at 1584 and 1639 cm^{-1}, they also have additional deviations in the lower frequency regions. Predicted bands at 1417 and 1169 cm^{-1} in the [CO_2^-,$N^{\omega\prime}$] and [N^-,CO,$N^{\omega\prime}$] structures, respectively, clearly result in intensities much greater than observed experimentally. Likewise, both structures predict additional bands in the 700–800 cm^{-1} region that are not observed experimentally.

Overall, the lowest energy [N,CO^-,$N^{\omega\prime}$]-*gggtgct* conformer is most consistent with the observed spectrum, and contributions from other conformers are not needed to account for the experimental observations. The possibility that [N,CO^-,$N^{\omega\prime}$]-*ggggggt* and [N,CO^-,$N^{\omega\prime}$]-*gcggttt* also contribute to the observed spectrum cannot be ruled out on a spectroscopic basis, but they are unlikely to be appreciably populated because of their relatively high energies. Indeed, an equilibrium distribution of conformers at 298 K would have a population of $\geq$99.95% of the [N,CO^-,$N^{\omega\prime}$]-*gggtgct* conformer at all levels of theory.

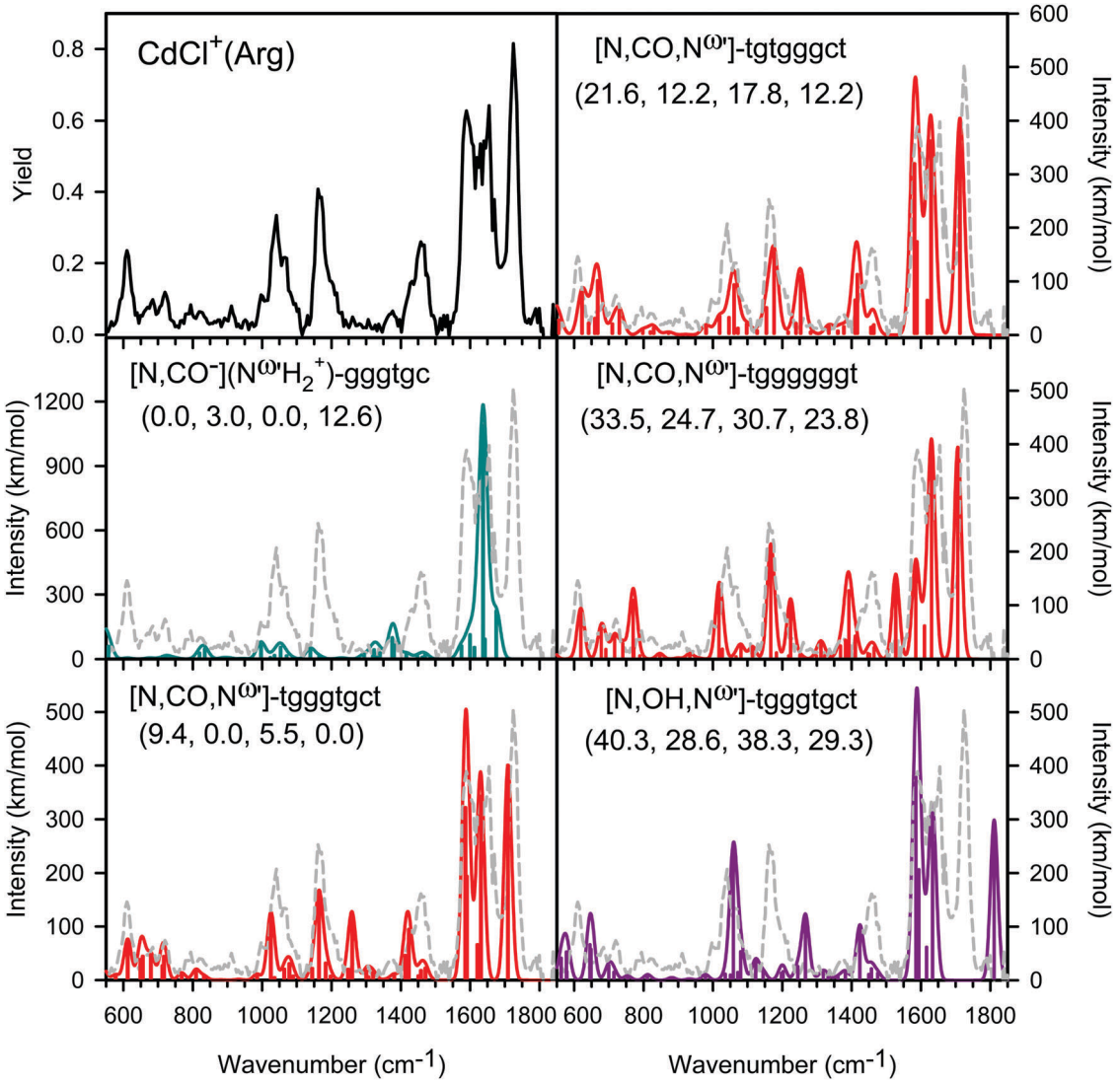

Fig. 4 Comparison of the $CdCl^+$(Arg) experimental IRMPD action spectrum (black, upper left; dashed grey, otherwise) with IR spectra calculated at the B3LYP/def2-TZVP level of theory for the lowest lying conformers (colored lines). Relative 0 K enthalpies (kJ mol^{-1}) are given at the B3LYP, B3LYP-GD3BJ, B3P86, and MP2(full) levels, respectively.

Comparison of theoretical and experimental IR spectra: $CdCl^+$(Arg)

The experimental IRMPD action spectrum of $CdCl^+$(Arg), Fig. 4, is characterized by dominant spectral features located at 610, 1041, 1161, 1457, 1589, 1655, and 1725 cm^{-1}. Also shown in Fig. 4 are the spectral comparisons between the five lowest energy $CdCl^+$(Arg) species and the experimental spectrum. An overview of the predicted vibrational frequencies and intensities of the five lowest energy conformers is given in the ESI,† Table S3. Interestingly, the $CdCl^+$(Arg) [N,CO^-]($N^{\omega\prime}H_2^+$)-*gggtgc* conformer (the B3LYP and B3P86 GS) does not reproduce the major spectral features well. One very intense band at 1637 cm^{-1} ($N^{\omega}H_2$ bend, CN^{ω} stretch, and CO stretch) is somewhat consistent with the experimental band at 1655 cm^{-1}, however it is red shifted by 18 cm^{-1} and is more than twice as intense as similar bands for other conformers. Thus, this species can only be a minor contributor to the observed spectrum. The minor experimental band at 1377 cm^{-1} is reproduced by this species (1375 and 1380 cm^{-1}, backbone CH_2 wag); however, all other experimental bands are more intense than those predicted by the [N,CO^-]($N^{\omega\prime}H_2^+$)-*gggtgc* calculated spectrum.

Instead, the $CdCl^+$(Arg) spectrum is reproduced reasonably well by the low-energy [N,CO,$N^{\omega\prime}$] tridentate structures, specifically [N,CO,$N^{\omega\prime}$]-*tgggtgct* (the B3LYP-GD3BJ and MP2(full) predicted GS). This conformer predicts an intense band at 1586 cm^{-1} ($N^{\omega\prime}CN^{\delta}$ asymmetric stretch and $CN^{\omega}H$ bend) and 1591 cm^{-1} (concerted guanidine bends), which reproduces the experimental 1589 cm^{-1} band quite well. Likewise, the experimental 1655 and 1725 cm^{-1} bands are reproduced reasonably well by this *tgggtgct* species, with predicted bands located at 1631 ($N^{\omega}H_2$ bend) and 1709 cm^{-1} (CO stretch and COH bend), although both bands are red-shifted slightly. The higher energy *tgtgggct* and *tgggggt* conformers reproduce these bands comparably, with predicted bands at 1582/1588, 1629, and 1713 cm^{-1} and 1527/1585, 1631, and 1706 cm^{-1}, respectively; however, the latter conformer fails to match the intensity of the experimental 1586 cm^{-1} band because the 1527 cm^{-1} peak no longer overlaps that of 1585 cm^{-1}, resulting in a separate peak not found in the experimental spectrum.

There is modest agreement between the various [N,CO,$N^{\omega\prime}$] species and the experimental spectrum in the lower frequency region below 1500 cm^{-1}, with the most favorable agreement again observed with the [N,CO,$N^{\omega\prime}$]-*tgggtgct* conformer. In particular, the experimental bands at 1041 and 1161 cm^{-1} are reproduced well by the predicted bands at 1025 cm^{-1} (backbone NH_2 wag) with 1079 cm^{-1} shoulder ($N^{\omega}H_2$ rock and CH_2 twists) and 1164 cm^{-1} (COH bend) of the *tgggtgct* conformer. The *tgtgggct* and *tggggggt* conformers reproduce the COH bend fairly well (1176 and 1167 cm^{-1}, respectively), but show discrepancies with experiment for the lower frequency NH_2 wag band (1061 and 1018 cm^{-1}, respectively). Likewise, the *tgggtgct* conformer reproduces the low intensity bands observed below 900 cm^{-1} reasonably well whereas the other two [N,CO,$N^{\omega\prime}$] conformers exhibit discrepancies mainly in intensities. None of these [N,CO,$N^{\omega\prime}$] species reproduce the experimental band at 1457 cm^{-1}. Instead, the *tgggtgct*, *tgtgggct*, and *tggggggt* conformers do have a band at 1423, 1417, and 1392 cm^{-1}, respectively, (antisymmetric combination of the $N^{\omega\prime}H$ and $N^{\delta}H$ bends), which appears to be blue shifted experimentally. Further, these conformers each predict a fairly intense band near 1250 cm^{-1} (largely $N^{\omega\prime}H$ bend with some symmetric $N^{\delta}H$ bend) that is not observed in the experimental spectrum. Here, a red-shift in this band would reproduce the shoulder to the blue of the 1161 cm^{-1} band observed experimentally. These two theoretical bands both involve primarily the $N^{\omega\prime}H$ bend combined with symmetric and antisymmetric motions of the $N^{\delta}H$ bend, such that it seems plausible these two bands could couple leading to the hypothesized blue and red shifts.

The [N,OH,$N^{\omega\prime}$] conformers, *tgggtgct* (shown in Fig. 4) and the very similar *tgtgggct* analogue, do not match the spectral features of the experimental spectrum well. The [N,OH,$N^{\omega\prime}$] conformers do not reproduce two key bands at 1161 and 1725 cm^{-1} because the hydroxyl group is bound directly to the metal ion preventing the coordinated CO stretch at 1725 cm^{-1} and shifting the COH bend associated with the band at 1161 cm^{-1}. This also results in an intense uncoordinated carbonyl stretch predicted at around 1810 cm^{-1}, for which there is no experimental evidence.

Ultimately, it is clear that the dominant ion present is [N,CO,$N^{\omega\prime}$]-*tgggtgct*, with contributions from the *tgtgggct* conformer possible on a spectroscopic basis, but unlikely given the much higher energy. However, although the [N,CO,$N^{\omega\prime}$]-*tgggtgct* conformer accounts for most of the experimental spectrum fairly well, observation of the minor band at 1377 cm^{-1} and the breadth of the band at 1655 cm^{-1} suggests that a minor population of the [N,CO^-]($N^{\omega\prime}H_2^+$)-*gggtgc* conformer may be contributing to the spectrum at 298 K. As shown in Fig. S5 (ESI†), a 85 : 15 [N,CO,$N^{\omega\prime}$]-*tgggtgct* : [N,CO^-]($N^{\omega\prime}H_2^+$)-*gggtgc* composite fit reproduces the experimental spectrum well, with an estimated uncertainty in the populations of ±5%. According to both B3LYP and B3P86 levels of theory, the [N,CO^-]($N^{\omega\prime}H_2^+$)-*gggtgc* conformer is the lowest energy structure, with an equilibrium distribution at 298 K of ≥98% and ≥94%, respectively. In contrast, the MP2(full) and B3LYP-GD3BJ levels predict that their [N,CO,$N^{\omega\prime}$]-*tgggtgct* GS would be populated at around 98% and ~70%, respectively, with the remaining population in [N,CO^-]($N^{\omega\prime}H_2^+$)-*gggtgc*. Here, it appears that the B3LYP and B3P86 functionals do not adequately predict the relative energetics for these heavy metal complexes, whereas perturbation-based MP2 performs better. B3LYP-GD3BJ generally reproduces MP2 values such that the good agreement at this level of theory is also reasonable. Apparently, the DFT methods without dispersion (which is designed to quantify hydrogen bonding more precisely) overvalue the $N^{\omega\prime}H^{\bullet}OC$ hydrogen bond in [N,CO^-]($N^{\omega\prime}H_2^+$)-*gggtgc* compared to the Cd–$N^{\omega\prime}$ interaction in [N,CO,$N^{\omega\prime}$]-*tgggtgct*.

Overall comparison

Because the Zn complex and the dominant Cd complex are assigned to tridentate structures in which the metal cation is bound to the backbone amine, carbonyl, and side-chain imine ($N^{\omega\prime}$), it is expected that they should share many spectral features. This comparison is shown in Fig. 5, with the main bands found in the $[Zn(Arg–H)]^+$ complex located at 632, 1067, 1150, 1246, 1424, 1584, 1639, and 1758 cm^{-1}, and those in the

$CdCl^+$(Arg) spectrum appearing at 610, 1041, 1161, 1457, 1589, 1655, and 1725 cm^{-1}. The highest wavenumber bands in both spectra correspond to carbonyl stretches and experience a sizable shift (33 cm^{-1}) from one another, which is explained by deprotonation of the carboxylic acid in the $[Zn(Arg–H)]^+$ species. This effect was previously noted in a study of histidine complexed with Zn^{2+} and Cd^{2+} where the recorded shift was about 40 cm^{-1}.[22] The bands at 1639 and 1655 cm^{-1} correspond primarily to the $N^{\omega}H_2$ bend (predicted to occur at 1632 and 1631 cm^{-1}, respectively) and are shifted in the experimental spectrum primarily because the Cd spectrum probably includes a contribution from the $[N,CO^-](N^{\omega\prime}H_2^+)$-*gggtgc* conformer at 1637 cm^{-1}. Bands at 1584 and 1589 cm^{-1} are comparable because they correspond to combinations of guanidine motions, and the same is true of the 1424 and 1457 cm^{-1} bands, where the shift in the latter band is discussed above. Likewise, as discussed above, the band at 1246 cm^{-1} in the Zn spectrum ($N^{\delta}H$ bend and CH_2 twists) appears red-shifted in the Cd spectrum, forming the shoulder near 1200 cm^{-1}. The bands near 1150 and 620 cm^{-1} are serendipitously similar as they correspond to a HNCH bend and out-of-plane $N^{\omega\prime}H$ bend in the Zn complex and in-plane and out-of-plane COH bends in the Cd spectrum. The bands near 1050 cm^{-1} both correspond to the backbone NH_2 wag. Some differences in the bands are expected because the larger Cd^{2+} ion should perturb the vibrations of the neutral Arg ligand less than the smaller Zn^{2+} ion perturbs the anionic deprotonated Arg ligand. Ultimately, it can be seen that the characteristic complexation and type of metal ion have major effects on the vibrational modes observed in the IRMPD action spectra.

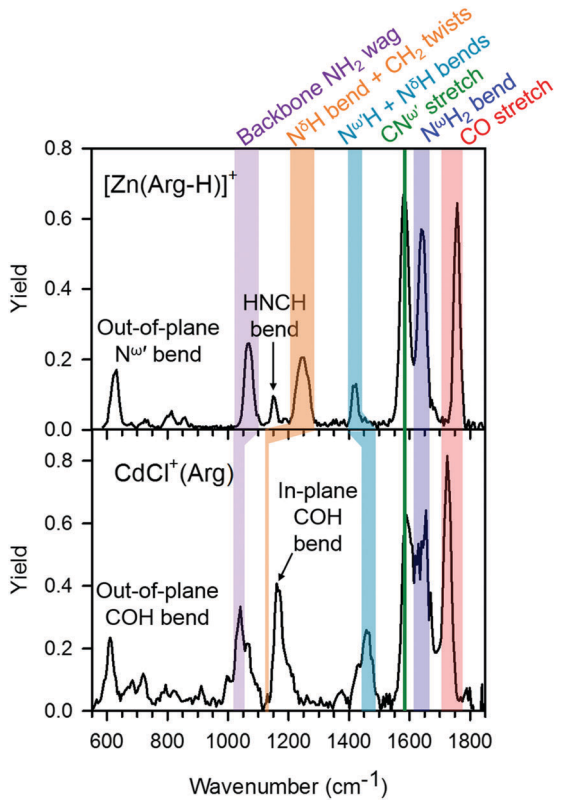

Fig. 5 Infrared multiple photon dissociation (IRMPD) action spectra of Zn^{2+} and Cd^{2+} complexed with Arg.

Comparison to IRMPD studies of Zn^{2+} and Cd^{2+} complexes with His

It is also of interest to compare the binding of Zn and Arg with that of Zn and His, one of the most common binding sites for Zn in biological systems. The binding of Zn to these two amino acids holds many similarities, including their nitrogen-based side-chain basicity and GS conformations, Fig. 6. In the gas-phase, the ground structures of Zn^{2+} binding to Arg and His both involve deprotonation at the carboxylic acid with $[N,CO^-,N^s]$ binding sites, where N^s is the $N^{\omega\prime}$ nitrogen in Arg and the N^{π} in His.[22] The Zn–O bond lengths are 1.88 and 1.90 Å, respectively, with Zn–N bonds lengths of 2.10 and 2.06 Å, respectively, and $r(Zn–N^{\omega\prime})$ = 1.92 Å compared to $r(Zn–N^{\pi})$ = 1.97 Å. The longer side-chain length in Arg probably allows more flexibility in the orientation of the side-chain binding site, allowing the shorter Zn–N^s bond distance. Indeed, such a shorter bond distance is consistent with the observation that Arg binds more strongly than His to Na^+, Cu^+, and Ag^+ in the gas phase.[46–49] In solution, protonation of the Arg side chain would disfavor metal binding. Probably the biggest difference between the two complexes is that the binding pocket in Arg is more planar ($\angle NZnON^{\omega\prime}$ = 155° and $\angle N^{\omega\prime}ZnO$ = 141°) compared to a more pyramidal orientation for His ($\angle NZnON^{\pi}$ = 100° and $\angle N^{\pi}ZnO$ = 112°). The latter orientation permits Zn to bind to a fourth ligand if present, whereas this would be hindered in the $[Zn(Arg–H)]^+$ complex. Notably, under similar electrospray ionization conditions, the $Zn^{2+}(His)_2$ complex (analogous to Zn^{2+}–His binding in Cys_2His_2 zinc finger domains) was formed in addition to a $[Zn(His–H)]^+(NH_3)$ complex (from which the $[Zn(His–H)]^+$ was formed), clearly different from the behavior observed in the current analysis where only $[Zn(Arg–H)]^+$ was observed. Comparison of these results suggests that Zn^{2+} binds preferentially to His (notably while keeping the His ligand intact in the larger complex) and is consistent with the characterization of prevalent histidine sites within zinc finger domains.

Similarly, the complexes formed between Cd^{2+} and Arg differ from those observed in our previous His study. There, both $Cd^{2+}(His)_2$ and $CdCl^+$(His) were formed in the source, whereas

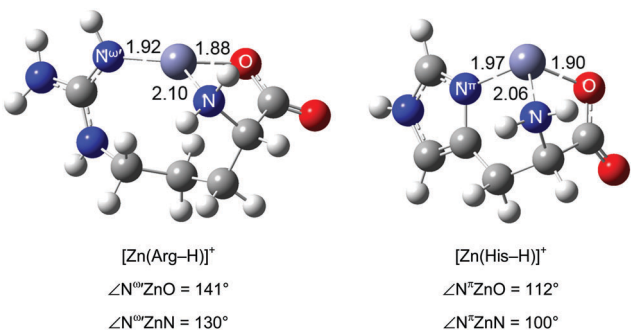

Fig. 6 Comparison of $[Zn(Arg–H)]^+$ and $[Zn(His–H)]^+$ ground structures.[22] Bond lengths (in Å) are indicated for select bonds.

only $CdCl^+$(Arg) was formed in the current study. In the His complexes, both $Cd^{2+}(His)_2$ and $CdCl^+$(His) GS conformations were characterized by [N,CO,N^{π}] binding motifs (see atom designations in Fig. 6), such that the ancillary Cl^- ligand does not greatly affect the binding motif. Similar results would be expected for the analogous comparison of $Cd^{2+}(Arg)_2$ and $CdCl^+$(Arg), although theoretical structures for $Cd^{2+}(Arg)_2$ were not explored because there was no experimental spectrum available for comparison.

Additional studies of interest could evaluate the effects of water solvation on the structural characteristics of these biologically important His and Arg systems, an analysis which could help correlate gas-phase results with observations made under physiological conditions. We expect that water solvation would only play a minor role in the complexation and stabilization of the current gas-phase systems (potentially acting mainly as a fourth ligand). Further, Cys_2His_2 zinc fingers have a hydrophobic core that allows the helix containing the two His residues to form,[1] such that water interactions at the active Zn binding site are probably not influential. Clearly, water solvation does play a role in the protein folding that leads to formation of the helix itself.

Conclusions

IRMPD action spectra of arginine complexed with Zn^{2+} and Cd^{2+} were measured between 600–1800 cm^{-1} and were compared to theoretical models calculated at the B3LYP/6-311+G(d,p) and B3LYP/def2-TZVP levels of theory, respectively. The theoretically predicted $[Zn(Arg–H)]^+$ GS, [N,CO^-,$N^{\omega\prime}$]-*gggtgct*, reproduces the experimental IRMPD action spectrum well. The contribution of other conformers is not needed to explain the observed spectrum and is unlikely because of their relatively high energies. Interestingly, the experimental $CdCl^+$(Arg) spectrum was determined to be a probable mixture of the two lowest-energy species, the bidentate [N,CO^-]($N^{\omega\prime}H_2^+$)-*gggtgc* and tridentate [N,CO,$N^{\omega\prime}$]-*tgggtgct* with the majority being the tridentate structure. This result is in better agreement with predictions of theory at the MP2(full) and B3LYP-GD3BJ levels than B3LYP and B3P86 levels. Analysis of the changes of vibrational modes and geometries of the arginine complexes provide knowledge of the relative binding strength of the metal ion within the complex. This in turn may be used to determine how biological systems react to differing metal centers. Shorter M–N, M–O, and M–$N^{\omega\prime}$ bond distances indicate that there is stronger Zn^{2+} binding to the complex compared to its Cd^{2+} congener. Overall, the information provided by this experiment contributes to our knowledge of metal–amino acid interactions and will help to further understand important biological systems.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Financial support for this work was provided by the National Science Foundation, Grants CHE-1664618 and OISE-1357887. We gratefully acknowledge the *Nederlandse Organisatie voor Wetenschappelijk Onderzoek* (NWO) for the support of the FELIX Laboratory as well as the staff at FELIX for their assistance. In addition, a grant of computer time from the Center of High Performance Computing at the University of Utah is greatly appreciated.

References

1 A. Klug, *Annu. Rev. Biochem.*, 2010, **79**, 213–231.
2 T. M. Hall, *Curr. Opin. Struct. Biol.*, 2005, **13**, 367–373.
3 S. Iuchi, *Cell. Mol. Life Sci.*, 2000, **58**, 625–635.
4 J. H. Laity, B. M. Lee and P. E. Wright, *Curr. Opin. Struct. Biol.*, 2001, **11**, 39–46.
5 R. J. Y. Simpson, E. D. Cram, R. Czolij, J. M. Matthews, M. Crossley and J. P. Mackay, *J. Biol. Chem.*, 2003, **278**, 28011–28018.
6 M. Imanishi, K. Matsumura, S. Tsuji, T. Nakaya, S. Negi, S. Futaki and Y. Sugiura, *Biochemistry*, 2012, **51**, 3342–3348.
7 S. Ling, W. Yu, Z. Huang, Z. Lin, M. Harañczyk and M. Gutowski, *J. Phys. Chem. A*, 2006, **110**, 12282–12291.
8 M. W. Forbes, M. F. Bush, N. C. Polfer, J. Oomens, R. C. Dunbar, E. R. Williams and R. A. Jockusch, *J. Phys. Chem. A*, 2007, **111**, 11759–11770.
9 M. F. Bush, J. T. O'Brien, J. S. Prell, R. J. Saykally and E. R. Williams, *J. Am. Chem. Soc.*, 2007, **129**, 1612–1622.
10 R. D. Shannon, *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.*, 1976, **32**, 751–767.
11 A. S. Lemoff, M. F. Bush, C.-C. Wu and E. R. Williams, *J. Am. Chem. Soc.*, 2005, **127**, 10276–10286.
12 M. F. Bush, J. Oomens, R. J. Saykally and E. R. Williams, *J. Phys. Chem. A*, 2008, **112**, 8578–8584.
13 N. C. Polfer, J. Oomens and R. C. Dunbar, *Phys. Chem. Chem. Phys.*, 2006, **8**, 2744–2751.
14 M. F. Bush, M. W. Forbes, R. A. Jockusch, J. Oomens, N. C. Polfer, R. J. Saykally and E. R. Williams, *J. Phys. Chem. A*, 2007, **111**, 7753–7760.
15 J. T. O'Brien, J. S. Prell, J. D. Steill, J. Oomens and E. R. Williams, *J. Phys. Chem. A*, 2008, **112**, 10823–10830.
16 P. B. Armentrout, M. T. Rodgers, J. Oomens and J. D. Steill, *J. Phys. Chem. A*, 2008, **112**, 2248–2257.
17 M. T. Rodgers, P. B. Armentrout, J. Oomens and J. D. Steill, *J. Phys. Chem. A*, 2008, **112**, 2258–2267.
18 M. Citir, E. M. S. Stennett, J. Oomens, J. D. Steill, M. T. Rodgers and P. B. Armentrout, *Int. J. Mass Spectrom.*, 2010, **297**, 9–17.
19 D. R. Carl, T. E. Cooper, J. Oomens, J. D. Steill and P. B. Armentrout, *Phys. Chem. Chem. Phys.*, 2010, **12**, 3384–3398.
20 M. Citir, C. S. Hinton, J. Oomens, J. D. Steill and P. B. Armentrout, *J. Phys. Chem. A*, 2012, **116**, 1532–1541.
21 N. C. Polfer, J. Oomens, D. T. Moore, G. von Helden, G. Meijer and R. C. Dunbar, *J. Am. Chem. Soc.*, 2006, **128**, 517–525.

22 T. E. Hofstetter, C. Howder, G. Berden, J. Oomens and P. B. Armentrout, *J. Phys. Chem. B*, 2011, **115**, 12648–12661.
23 A. Gholami and T. D. Fridgen, *J. Phys. Chem. B*, 2013, **117**, 8447–8456.
24 R. A. Coates, C. P. McNary, G. C. Boles, G. Berden, J. Oomens and P. B. Armentrout, *Phys. Chem. Chem. Phys.*, 2015, **17**, 25799–25808.
25 G. C. Boles, R. A. Coates, G. Berden, J. Oomens and P. B. Armentrout, *J. Phys. Chem. B*, 2015, **119**, 11607–11617.
26 R. A. Coates, G. C. Boles, C. P. McNary, G. Berden, J. Oomens and P. B. Armentrout, *Phys. Chem. Chem. Phys.*, 2016, **18**, 22434–22445.
27 G. C. Boles, R. A. Coates, G. Berden, J. Oomens and P. B. Armentrout, *J. Phys. Chem. B*, 2016, **120**, 12486–12500.
28 G. C. Boles, C. J. Owen, G. Berden, J. Oomens and P. B. Armentrout, *Phys. Chem. Chem. Phys.*, 2017, **19**, 12394–12406.
29 G. C. Boles, R. L. Hightower, R. A. Coates, C. P. McNary, G. Berden, J. Oomens and P. B. Armentrout, *J. Phys. Chem. B*, 2018, **122**, 3836–3853.
30 C. J. Owen, G. C. Boles, G. Berden, J. Oomens and P. B. Armentrout, *Eur. J. Mass Spectrom.*, 2018, DOI: 10.1177/1469066718792902.
31 D. Oepts, A. F. G. van der Meer and P. W. van Amersfoort, *Infrared Phys. Technol.*, 1995, **36**, 297–308.
32 J. J. Valle, J. R. Eyler, J. Oomens, D. T. Moore, A. F. G. van der Meer, G. von Heldon, G. Meijer, C. L. Hendrickson, A. G. Marshall and G. T. Blakney, *Rev. Sci. Instrum.*, 2005, **76**, 023103.
33 N. C. Polfer and J. Oomens, *Phys. Chem. Chem. Phys.*, 2007, **9**, 3804–3817.
34 S. Guan and A. G. Marshall, *Int. J. Mass Spectrom. Ion Processes*, 1996, **157–158**, 5–37.
35 A. G. Marshall, T. C. L. Wang and T. L. Ricca, *J. Am. Chem. Soc.*, 1985, **107**, 7893–7897.
36 R. A. Jockusch, W. D. Price and E. R. Williams, *J. Phys. Chem. A*, 1999, **103**, 9266–9274.
37 M. J. Frisch, G. 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, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Revision D.01*, Gaussian, Inc., Wallingford, CT, USA, 2009.
38 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
39 D. Feller, *J. Comput. Chem.*, 1996, **17**, 1571–1586.
40 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.
41 N. C. Polfer, *Chem. Soc. Rev.*, 2011, **40**, 2211–2221.
42 M. K. Kesharwani, B. Brauer and J. M. L. Martin, *J. Phys. Chem. A*, 2015, **119**, 1701–1714.
43 M. Berglund and M. E. Wieser, *Pure Appl. Chem.*, 2011, **83**, 397–410.
44 H. Lavanant and Y. Hoppilliard, *J. Mass Spectrom.*, 1997, **32**, 1037–1049.
45 H. Lavanant and Y. Hoppilliard, *Eur. Mass Spectrom.*, 1999, **5**, 41–50.
46 P. Wang, G. Ohanessian and C. Wesdemiotis, *Int. J. Mass Spectrom.*, 2008, **269**, 34–45.
47 B. A. Cerda and C. Wesdemiotis, *J. Am. Chem. Soc.*, 1995, **117**, 9734–9739.
48 V. W.-M. Lee, H. Li, T.-C. Lau, R. Guevremont and K. W. M. Siu, *J. Am. Soc. Mass Spectrom*, 1998, **9**, 760–766.
49 P. B. Armentrout and M. T. Rodgers, *Adv. Mass Spectrom.*, 2013, **19**, 47–58.