

# PCCP

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

## ARTICLE

# The Structures and Properties of Proton- and Alkali-bound Cysteine Dimers

Cite this: DOI: 10.1039/x0xx00000x

Christian Ieritano,<sup>1</sup> Patrick J. J. Carr,<sup>1</sup> Moaraj Hasan,<sup>1</sup> Michael Burt,<sup>1</sup> Rick A. Marta,<sup>1</sup> Vincent Steinmetz,<sup>2</sup> Eric Fillion,<sup>1</sup> Terrance B. McMahan,<sup>1</sup> W. Scott Hopkins<sup>1,\*</sup>

Received 01st December 2015,  
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

The proton-, lithium-, and sodium-bound cysteine dimers have been investigated in a joint computational and experimental infrared multiple photon dissociation (IRMPD) study. IRMPD spectra in the 1000–2000  $\text{cm}^{-1}$  region show that protonation is localized on an amine group, and that intermolecular hydrogen bonding occurs between the protonated amine and the carbonyl oxygen of the neutral Cys moiety. Alkali-bound dimers adopt structures reminiscent of those observed for the monomeric  $\text{Cys}\cdot\text{Li}^+$  and  $\text{Cys}\cdot\text{Na}^+$  species. Calculations of the heavier  $\text{Cys}_2\cdot\text{M}^+$  ( $\text{M} = \text{K}, \text{Rb}$  or  $\text{Cs}$ ) species suggest that these are significantly less strongly bound than the lighter ( $\text{M} = \text{H}, \text{Li}$ , or  $\text{Na}$ ) dimers.

## Introduction

Cysteine (Cys) plays an important structural role in many proteins owing to the fact that its thiol side chain can coordinate with metal cations or oxidize to form disulphide linkages with other Cys residues.<sup>1–6</sup> Typically, Cys/ion interactions in proteins involve multiply charged metal centers (e.g., Zn in zinc finger structures),<sup>1–3</sup> which raises the question as to why trace metals compete more effectively for Cys-containing protein binding sites than do the more common singly charged alkali metal cations.<sup>7, 8</sup> It is therefore of interest to determine the structures and properties of Cys complexes with a variety of metal cations, since this might shed light on metal ion binding motifs in larger, biologically relevant structures.

To examine the binding between the alkali metal cations and Cys, Armentrout and coworkers investigated pairwise interactions using guided ion beam mass spectrometry.<sup>9</sup> This work was followed with an infrared multiple photon dissociation (IRMPD) spectroscopic study, which showed that alkali metal cations adopt charge-solvated tridentate structures wherein the metal center is bound to the amino N, carbonyl O, and thiol S atoms.<sup>7</sup> For the heavier alkali metal species ( $\text{K}^+$ ,  $\text{Rb}^+$  or  $\text{Cs}^+$ ) additional isomeric structures were also thought to be observed – most notably in the case of  $\text{Cys}\cdot\text{Cs}^+$  where the tridentate conformation is slightly higher in energy than the calculated (bidentate) global minimum structure. Comparative investigations of transition metal Cys complexes have been challenging owing to the fact that d-block metals apparently catalyze oxidation of Cys to the disulphide-linked dimer cystine.<sup>5</sup> Nevertheless, the Oomens and Armentrout groups have recently persevered in this regard, collaborating on a structural characterization of gas-phase Cys monomeric complexes with zinc and cadmium.<sup>8</sup>

The oxidation of Cys to cystine is a dominant process in electrospray ionization (ESI).<sup>5</sup> Steill et al. reported that positive mode ESI of a 40/60 vol% water/methanol Cys solution yields predominantly protonated cystine,  $(\text{CySSCy})\cdot\text{H}^+$ , with minor

production (ca. 18–20%) of  $\text{Cys}\cdot\text{H}^+$  and generation of trace amounts of the protonated Cys dimer,  $(\text{Cys})_2\cdot\text{H}^+$ .<sup>5</sup> The process by which cystine forms is thought to occur via a radical pathway.<sup>4</sup> Owing to the instability of the Cys radicals, dimerization to form cystine is expected to be rapid. To date, this has been an impediment to the study of non-covalently bound Cys dimers. Here, we report on a combined computational and experimental study of proton- and alkali-bound Cys dimers. Experimental outcomes were supported by detailed basin-hopping molecular dynamics searches of the cluster potential energy surfaces (PESs) and high-level DFT calculations.

## Experimental

IRMPD spectra of proton-, lithium-, and sodium-bound cysteine dimers were recorded at the Centre Laser Infrarouge d'Orsay (CLIO) free electron laser (FEL) facility at the University of Paris XI.<sup>10, 11</sup> Details of the experimental apparatus have been published previously.<sup>12–14</sup> Proton-, lithium-, and sodium-bound dimers of cysteine were produced by electrospray ionization (ESI) of ca. 100  $\mu\text{mol}\cdot\text{L}^{-1}$  methanol/water (50/50 vol%) solutions containing L-cysteine, the alkali chloride salt of interest, and formic acid (ca. 5  $\mu\text{mol}\cdot\text{L}^{-1}$ ). The ESI conditions were such that the potassium-, rubidium-, and cesium-bound dimers of cysteine could not be produced. The cationized monomers of all species were, however, formed in relatively high abundance. Nascent clusters produced by ESI were transferred to a Bruker Esquire 3000+ ion trap mass spectrometer, where they were mass-selected and subsequently irradiated by the tunable output of the FEL over the 1000 – 2000  $\text{cm}^{-1}$  range. Vibrational spectra were generated by recording the fragmentation efficiency of the cysteine clusters as a function of FEL wavenumber.

## Computational Methods

A custom-written basin-hopping (BH) code that is interfaced with Gaussian 09 was used to map the cysteine cluster potential energy surfaces (PESs).<sup>12, 15-18</sup> The clusters were modelled using the AMBER molecular mechanics force field, which incorporated atomic partial charges for cysteine as calculated at the B3LYP/6-31+G(d,p) level of theory using the CHelpG partition scheme.<sup>19,21</sup> Each of the cationized monomers (*viz.* Cys•M<sup>+</sup>; M = H, Li – Cs) and the proton-/alkali-bound dimers (*viz.* (Cys)<sub>2</sub>•M<sup>+</sup>) were studied computationally. For each random structural perturbation in the BH search, the charge-carrying species (*viz.* H<sup>+</sup>, Li<sup>+</sup>, Na<sup>+</sup>, K<sup>+</sup>, Rb<sup>+</sup>, or Cs<sup>+</sup>) was given a random translational step of  $-0.7 \text{ \AA} \leq \eta \leq 0.7 \text{ \AA}$  in each of the X, Y, and Z directions. For the Cys•M<sup>+</sup> species each of the dihedral angles in cysteine was given a random rotation of  $-20^\circ \leq \phi \leq 20^\circ$ . The same random dihedral angle rotation was applied for the (Cys)<sub>2</sub>•M<sup>+</sup> clusters. In this case, however, one of the cysteine molecules was given a random translational step of  $-0.7 \text{ \AA} \leq \eta \leq 0.7 \text{ \AA}$  in each of the X, Y, and Z directions as well as random rotation of  $-20^\circ \leq \theta \leq 20^\circ$  about its x, y, and z body-fixed axes.

In total, *ca.* 40,000 structures were sampled by the BH algorithm for each cluster. The BH routine identified more than 500 unique conformations for each (Cys)<sub>2</sub>•M<sup>+</sup> cluster. The 50 lowest energy isomers from each search were pre-optimized at the HF/6-31+G(d,p) level of theory, and unique isomers from this step were then carried forward for treatment at the B3LYP/6-311++G(d,p) level of theory. The LANL2DZ basis set and pseudopotential were employed for Rb and Cs calculations. To ensure that each isomer was a local minimum on the PES, normal mode analyses were conducted. This also served to calculate the harmonic vibrational spectrum for each structure. For selected isomers, anharmonically corrected vibrational frequencies were also calculated. DFT results and cluster XYZ coordinates are provided in the supplementary information.

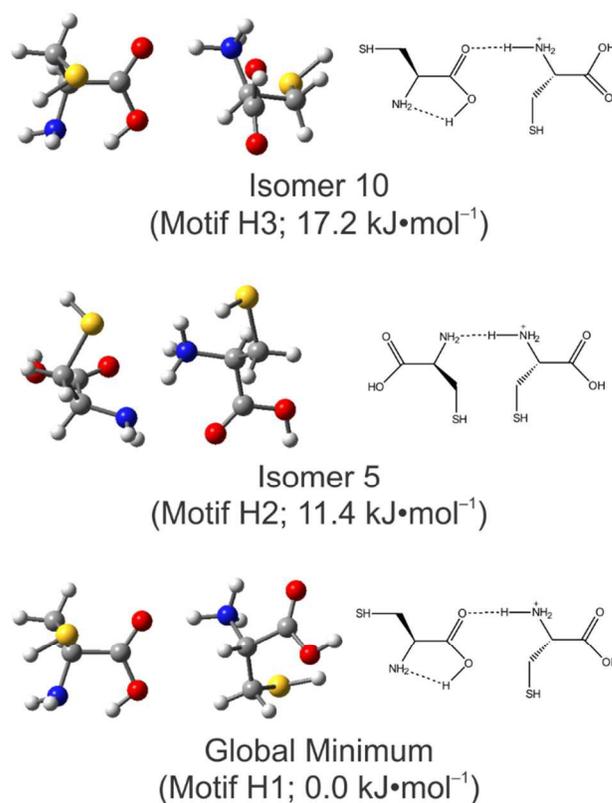
## Results and Discussion

### The Proton-bound Dimer, (Cys)<sub>2</sub>•H<sup>+</sup>

Three low-energy binding motifs for (Cys)<sub>2</sub>•H<sup>+</sup> were identified in our computational study. The lowest energy isomer for each of these motifs are shown in Figure 1. Isomers are numbered in order of increasing zero-point corrected electronic energy (*viz.* isomers 2 – 4 also adopt binding motif H1). Calculated thermodynamic corrections are provided in the supplementary information. In all cases, protonation occurs on the amine site. For motif H1, the protonated amine participates in an intermolecular hydrogen bond with the carbonyl oxygen of the neutral Cys moiety, while in the case of motif H2, the proton is shared between the amine groups of the two Cys molecules. Motif H3 differs from motif H1 by a 180° rotation of the NCCO dihedral angle of the amino acid functional group. Note that motifs H1 and H3 both exhibit intramolecular hydrogen bonding between the hydroxyl and amine groups of the neutral Cys moiety.

The IRMPD spectrum for (Cys)<sub>2</sub>•H<sup>+</sup> is shown in Figure 2 along with the calculated harmonic spectra of isomers 1 (global minimum), 5, and 10 (see Figure 1). Clearly, there is excellent agreement between the B3LYP predicted spectrum for the global minimum structure and the vibrational spectrum that is observed experimentally. Based on the calculations, the vibrational peak at 1752 cm<sup>-1</sup> is assigned to the free C=O stretch in the protonated Cys moiety, whereas the carbonyl group that is participating in the hydrogen bonding interaction with the NH<sub>3</sub><sup>+</sup> group (see Figure 1) exhibits an absorption at 1701 cm<sup>-1</sup>. The two intense peaks at 1381 cm<sup>-1</sup> and 1149 cm<sup>-1</sup> are associated with COH bending motions in the neutral and protonated Cys moieties, respectively. However, the band at 1149 cm<sup>-1</sup> simultaneously excites large amplitude rocking motion of the NH<sub>3</sub><sup>+</sup> group. The NH<sub>3</sub><sup>+</sup> group also exhibits an umbrella

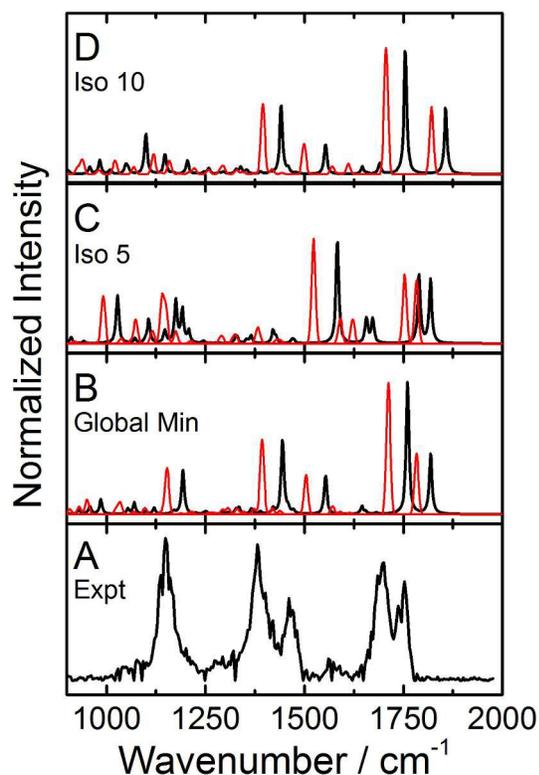
motion, which is observed at 1461 cm<sup>-1</sup>. These latter two transitions, in particular, provide direct evidence that protonation is localized on one of the amine groups. This conclusion could be confirmed by the observation of an intense band predicted to be at 2836 cm<sup>-1</sup>, which is associated with the H<sub>2</sub>N<sup>+</sup>–H•••O=C stretching mode of the shared proton. However, this band falls outside of the CLIO scan range and is likely to be significantly broadened owing to the anharmonic nature of the shared proton vibrational motion.<sup>5, 22, 23</sup> The red traces in Figure 2 show the calculated spectra following anharmonic correction. The relatively small spectral shift to lower wavenumber (when compared with the harmonic spectrum; ~45 cm<sup>-1</sup>) indicates that the observed transitions are all associated with relatively harmonic vibrational motions. This is also true for the alkali-bound dimers (*vide infra*). When considering the five most intense computed harmonic *versus* anharmonic transitions for (Cys)<sub>2</sub>•H<sup>+</sup>, an average scaling factor of 0.9710 is calculated, which is in excellent agreement with the commonly employed literature value of 0.9679.<sup>24</sup>



**Figure 1.** The lowest energy structures for the three lowest energy structural motifs of (Cys)<sub>2</sub>•H<sup>+</sup> as calculated at the B3LYP/6-311++G(d,p) level of theory. Electronic energies are zero-point corrected.

The observed IRMPD peak widths of *ca.* 50 cm<sup>-1</sup> are larger than the typical CLIO bandwidth of 25 cm<sup>-1</sup>,<sup>10</sup> thus suggesting that there are at least two isomers contributing to the (Cys)<sub>2</sub>•H<sup>+</sup> spectrum. The fact that these additional isomers yield peaks that are very near to those of the predicted global minimum isomer indicates that the observed higher energy isomers are structurally very similar to the global minimum. For example, the global minimum and isomers 2–4 all share the same binding motif, but differ in conformation by dihedral rotations about single bonds of the carbon side-chains in the Cys moieties. Consequently, their calculated spectra are also very similar to one another (see supplementary information). In isomer 5, on the other hand, the proton is shared between the two amine groups and there is no intramolecular hydrogen bonding within the

Cys moieties. This results in a significantly different cluster structure and, therefore, a considerably different vibrational spectrum than that of the ground state isomer (see Figure 2). Binding motif H3 is, however, very similar to motif H1. Consequently, the calculated spectra for isomer 10 (motif H3) and the global minimum (motif H1) are very similar. There are, however, subtle differences between the calculated spectra for isomers 1 and 10; the splitting between the free and hydrogen-bound carbonyl peaks is larger for isomer 10 compared to the global minimum, and there is little intensity in the  $1150\text{ cm}^{-1}$  region of the anharmonically corrected spectrum of isomer 10. While it is tempting, based on these calculated differences, to conclude that isomer 10 is not present in the sample, it is not uncommon in IRMPD spectra to observe transition intensities that are very different from those determined by electronic structure calculations.<sup>12, 25–29</sup> Given the calculated difference in relative energy between isomers 1 and 10, though, we would expect that motif H3 is (at best) a minor contributor to the observed vibrational spectrum.

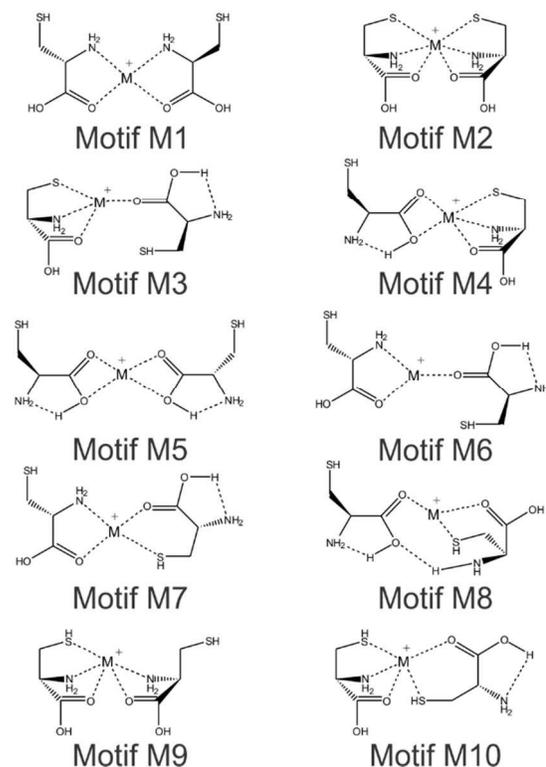


**Figure 2.** The (A) experimental IRMPD spectrum of  $(\text{Cys})_2\cdot\text{H}^+$  and the calculated harmonic spectra of the (B) global minimum, (C) isomer 5 ( $11.4\text{ kJ}\cdot\text{mol}^{-1}$ ), and (D) isomer 10 ( $17.2\text{ kJ}\cdot\text{mol}^{-1}$ ) structures. Overlaid red traces show the calculated anharmonic spectra. Spectra are broadened with a  $4\text{ cm}^{-1}$  Gaussian linewidth. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory.

### The Alkali-bound Dimers, $(\text{Cys})_2\cdot\text{M}^+$ ( $\text{M} = \text{Li}, \text{Na}, \text{K}, \text{Rb}$ or $\text{Cs}$ )

Our search of the PESs for  $(\text{Cys})_2\cdot\text{M}^+$  ( $\text{M} = \text{Li}, \text{Na}, \text{K}, \text{Rb}$  or  $\text{Cs}$ ) identified several low energy binding motifs. The ten lowest energy alkali cation binding motifs are shown in Figure 3. Note that the relative energies of these motifs change from one alkali metal-bound dimer to another (*i.e.*, the motif energy ordering for  $(\text{Cys})_2\cdot\text{Li}^+$  is different from that of  $(\text{Cys})_2\cdot\text{Na}^+$ ). In principle, one could adopt the nomenclature for the binding

motifs of  $\text{Cys}\cdot\text{M}^+$  species to describe the alkali-bound dimers.<sup>7</sup> For example, motif M1 exhibits bidentate [N,CO] binding for both Cys moieties under this labelling scheme. However, this nomenclature becomes somewhat cumbersome for larger clusters, especially when including the *cis/gauche/trans* (*i.e.*, *syn/gauche/anti*) descriptions for the side-chain dihedral orientations. For this reason, and because our experiments seem not to be particularly sensitive to side-chain conformational differences for a given binding motif, we will refer to the binding motifs shown in Figure 3 in our subsequent discussion.



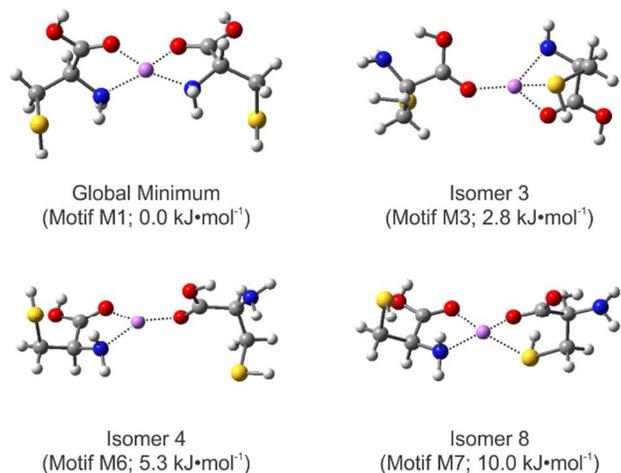
**Figure 3.** The low energy binding motifs of alkali-bound cysteine dimers,  $(\text{Cys})_2\cdot\text{M}^+$  ( $\text{M} = \text{Li}, \text{Na}, \text{K}, \text{Rb}$  or  $\text{Cs}$ ).

As one might expect, given the observed binding motifs for the  $\text{Cys}\cdot\text{M}^+$  monomers, bidentate and tridentate Cys binding motifs are common for the alkali-bound dimers.<sup>7</sup> Intramolecular  $\text{OH}\cdots\text{NH}_2$  hydrogen bonding also commonly occurs, and it is seen in all cases of monodentate binding. Motifs that exhibit intermolecular hydrogen bonding are predicted to occur at slightly higher energies, but there is no evidence for these structures (or zwitterionic structures) experimentally (*vide infra*).

#### i. The Lithium-bound Dimer, $(\text{Cys})_2\cdot\text{Li}^+$

Figure 4 shows the four lowest energy binding motifs for the lithium-bound Cys dimer. Isomers are numbered in order of increasing energy relative to the global minimum (*i.e.*, isomer 1). In other words, there were eight isomers associated with four different binding motifs calculated to lie below  $10\text{ kJ}\cdot\text{mol}^{-1}$  relative to the calculated global minimum. Note that three of the four lowest energy  $(\text{Cys})_2\cdot\text{Li}^+$  binding motifs exhibit (distorted) tetrahedral coordination to the central lithium

cation, while isomer 4 (motif M6) exhibits a trigonal planar geometry about the  $\text{Li}^+$  centre. This is consistent with the small, non-polarizable nature of  $\text{Li}^+$ . In their study of monomeric  $\text{Cys}\cdot\text{M}^+$  ( $\text{M} = \text{Li}, \text{Na}, \text{K}, \text{Rb}$  or  $\text{Cs}$ ) species, Citir *et al.* showed that  $\text{Cys}\cdot\text{Li}^+$  adopts tridentate coordination via the N, S, and carbonyl O. Interestingly, calculations suggest that  $(\text{Cys})_2\cdot\text{Li}^+$  should exhibit (for the most part) bidentate coordination between Cys and the central lithium cation. Only motif M3 (e.g., isomer 3) exhibits a tridentate coordination to one of the Cys moieties (monodentate coordination to the other Cys).

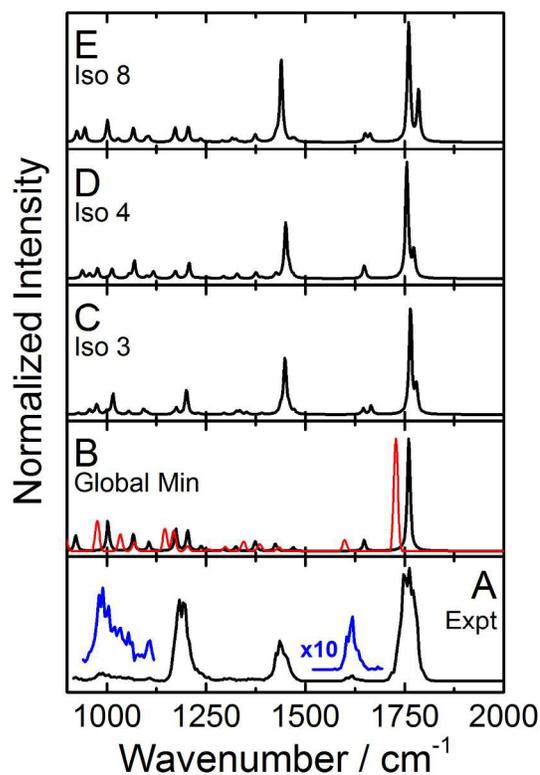


**Figure 4.** The lowest energy structures for the four lowest energy structural motifs of  $(\text{Cys})_2\cdot\text{Li}^+$  as calculated at the B3LYP/6-311++G(d,p) level of theory. Electronic energies are zero-point corrected.

Figure 5A shows the experimental IRMPD spectrum for  $(\text{Cys})_2\cdot\text{Li}^+$ . Three major peaks are observed:  $1187 \text{ cm}^{-1}$  (free COH bend),  $1436 \text{ cm}^{-1}$  (H-bonded COH bend), and  $1758 \text{ cm}^{-1}$  (C=O stretch). Weak bands observed at  $1625 \text{ cm}^{-1}$  and  $1000 \text{ cm}^{-1}$  are assigned to  $\text{NH}_2$  scissoring and combined  $\text{NH}_2$  wagging / CSH bending modes, respectively. The fact that a single (relatively broad) peak is observed in the carbonyl stretching region indicates that both C=O groups are in similar chemical environments. This observation accords well with all four of the binding motifs shown in Figure 4, where in each case both carbonyl groups are coordinated to the central lithium cation. Consequently, the calculated harmonic spectra for isomers 1, 3, 4, and 8 (Figures 5B-E) all exhibit vibrational transitions in the *ca.*  $1750 \text{ cm}^{-1}$  region. Again, peak widths of *ca.*  $50 \text{ cm}^{-1}$  suggest that (at least) two isomers with very similar structures contribute to the observed spectrum. In comparing the experimental and calculated spectra, we find that the spectrum for isomer 3 provides the best match. This structure may be viewed as a  $[\text{N},\text{CO},\text{S}]$  lithiated monomer (as identified by Citir *et al.*)<sup>7</sup> which is bound to a neutral Cys moiety via its carbonyl oxygen. Note that isomer 3 exhibits the lowest standard Gibbs energy of the four isomers shown. Given the similarity amongst the calculated spectra for the four low energy binding motifs, we cannot discount any of these four structures from being present in the  $(\text{Cys})_2\cdot\text{Li}^+$  ensemble.

The onset of the fifth structural motif occurs at  $23.9 \text{ kJ}\cdot\text{mol}^{-1}$  above the calculated global minimum (isomer 17;  $25.6$

$\text{kJ}\cdot\text{mol}^{-1}$  when accounting for thermal correction at  $298 \text{ K}$ ). This binding motif (not shown in Figure 3; see supplementary information) involves monodentate coordination of both carbonyl oxygen atoms to the  $\text{Li}^+$  centre and two intramolecular  $\text{OH}\cdots\text{NH}_2$  hydrogen bonds (one within each Cys moiety). Consequently, this motif does not exhibit a peak in the  $1190 \text{ cm}^{-1}$  where free COH bends are observed. This species does, however, exhibit vibrational peaks that correspond with the other bands observed experimentally, so it is possible that this motif, too, is present in the ensemble. Given the relative energy of this motif, though, we think it unlikely to be present. Especially so since the sixth structural motif for  $(\text{Cys})_2\cdot\text{Li}^+$ , which is a zwitterionic version of isomer 4, occurs at  $24.6 \text{ kJ}\cdot\text{mol}^{-1}$  above the calculated global minimum and it can be definitively discounted based on its calculated spectrum.



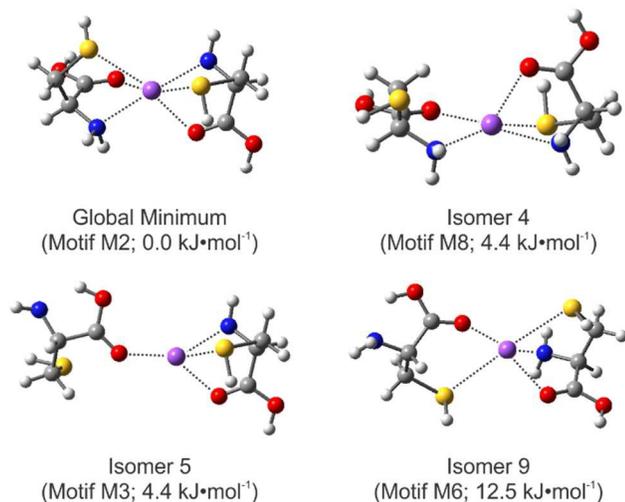
**Figure 5.** The (A) experimental IRMPD spectrum of  $(\text{Cys})_2\cdot\text{Li}^+$  and the calculated harmonic spectra of the (B) global minimum, (C) isomer 3 ( $2.8 \text{ kJ}\cdot\text{mol}^{-1}$ ), (D) isomer 4 ( $5.3 \text{ kJ}\cdot\text{mol}^{-1}$ ), and (E) isomer 8 ( $10.0 \text{ kJ}\cdot\text{mol}^{-1}$ ) structures. The overlaid red trace shows the calculated anharmonic spectrum for the global minimum. Spectra are broadened with a  $4 \text{ cm}^{-1}$  Gaussian linewidth. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory.

## ii. The Sodium-bound Dimer, $(\text{Cys})_2\cdot\text{Na}^+$

The four lowest energy binding motifs of the sodium-bound Cys dimer are shown in Figure 6. Isomers are numbered in order of increasing energy relative to the calculated global minimum. In general,  $(\text{Cys})_2\cdot\text{Na}^+$  favours higher coordination to the central alkali cation than does the lithiated analogue; in all four of the lowest energy binding motifs, at least one Cys moiety exhibits tridentate coordination. Note also that tridentate binding occurs via the amine

N, carbonyl O, and thiol S in all four of these species, which accords with the observed binding for the sodiated monomer.<sup>7</sup> The four binding motifs shown in Figure 6 all exhibit different geometries about the metal centre. The Global minimum structure (motif M2) exhibits a distorted octahedral geometry about the Na<sup>+</sup> centre, whereas isomer 4 (motif M8) has a distorted square based pyramidal geometry (viz. half of the octahedron). Isomer 5 exhibits a similar tetrahedral binding motif (motif M3) as was observed for (Cys)<sub>2</sub>•Li<sup>+</sup>, and isomer 9 (motif M10) displays a distorted trigonal bipyramidal geometry with respect to the Na cation.

The experimentally observed IRMPD spectrum of (Cys)<sub>2</sub>•Na<sup>+</sup> is plotted in Figure 7A. By comparing the observed spectrum with those that are calculated for isomers 1, 4, 5, and 9, we are able to assign the observed vibrational peaks: *ca.* 1750 cm<sup>-1</sup> (C=O stretch), *ca.* 1625 cm<sup>-1</sup> (NH<sub>2</sub> scissor), *ca.* 1450 cm<sup>-1</sup> (H-bonded COH bend), *ca.* 1200 cm<sup>-1</sup> (free COH bend), *ca.* 1120 cm<sup>-1</sup> (C–N stretch / carbon chain vibrations), *ca.* 980 cm<sup>-1</sup> (NH<sub>2</sub> wag / CSH bend). Note that these are the same peaks as were observed for (Cys)<sub>2</sub>•Li<sup>+</sup>; the spectrum for (Cys)<sub>2</sub>•Na<sup>+</sup> differs (predominantly) from that of the lithiated analogue in relative peak intensities and widths. For (Cys)<sub>2</sub>•Na<sup>+</sup>, peak FWHM ≈ 80 cm<sup>-1</sup> and (by inspection) the observed spectrum appears to be a convolution of several isomers. Interestingly, as was the case with (Cys)<sub>2</sub>•Li<sup>+</sup>, it is motif M3 that provides the best match with the experimental (Cys)<sub>2</sub>•Na<sup>+</sup> IRMPD spectrum (see Figure 7D). Furthermore, similar to (Cys)<sub>2</sub>•Li<sup>+</sup>, it was this structure which yielded the lowest standard Gibbs energy of the isomers shown in Figure 6. Motif M10 (*e.g.*, isomer 9), which differs from motif M3 by an additional S•••Na<sup>+</sup> interaction, also yields a reasonable match with experiment.



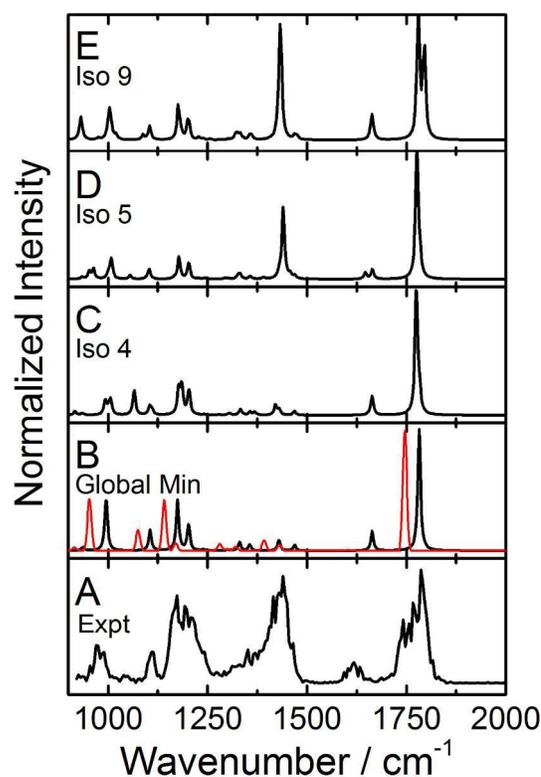
**Figure 6.** The lowest energy structures for the four lowest energy structural motifs of (Cys)<sub>2</sub>•Na<sup>+</sup> as calculated at the B3LYP/6-311++G(d,p) level of theory. Electronic energies are zero-point corrected.

In total, there are ten (Cys)<sub>2</sub>•Na<sup>+</sup> binding motifs below 22 kJ•mol<sup>-1</sup>. The first motif to higher energy than those shown in Figure 6 occurs at 16.2 kJ•mol<sup>-1</sup>. This structure (isomer 11) is a variant on motif M3 (*e.g.*, (Cys)<sub>2</sub>•Na<sup>+</sup> isomer 5) wherein the thiol group of the monodentate ligand is participating in intramolecular hydrogen bonding with the carbonyl oxygen. Consequently, the calculated IR spectrum for isomer 11 is also a reasonable match with

experiment. Cluster binding motifs to higher energy of isomer 11 do not yield vibrational spectra that match well with experiment.

### iii. Heavier Alkali-bound Dimers, (Cys)<sub>2</sub>•M<sup>+</sup> (M = K, Rb or Cs)

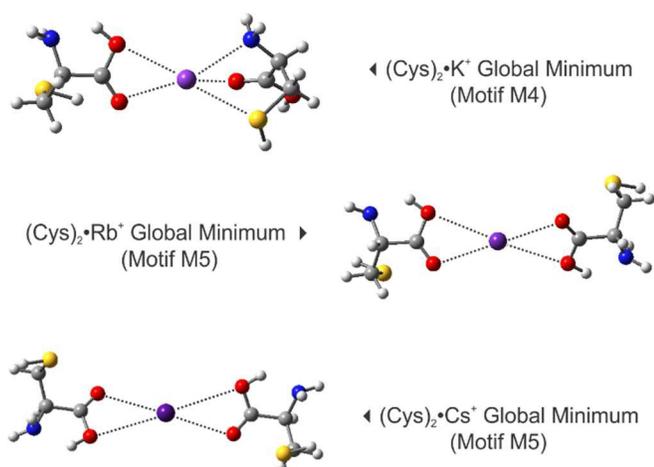
Although the (Cys)<sub>2</sub>•M<sup>+</sup> (M = K, Rb or Cs) clusters could not be produced for experimental study, these species were investigated computationally. The calculated global minima of (Cys)<sub>2</sub>•M<sup>+</sup> (M = K, Rb or Cs) are shown in Figure 8. (Cys)<sub>2</sub>•K<sup>+</sup> adopts binding motif M4, which is structurally related to motif M3, differing only by an additional HO•••M<sup>+</sup> interaction (*i.e.*, bi- instead of monodentate coordination of the second Cys). The result is a distorted square based pyramidal geometry about the K<sup>+</sup> centre. This is consistent with the Cys•K<sup>+</sup> monomer spectrum, which is a convolution of isomers associated with the [N,CO,S] and [COOH] binding motifs.<sup>7</sup> Both (Cys)<sub>2</sub>•Rb<sup>+</sup> and (Cys)<sub>2</sub>•Cs<sup>+</sup> are predicted to adopt binding motif M5, wherein both Cys moieties undergo bidentate coordination via the carboxylic acid oxygen atoms, and both exhibit intramolecular OH•••NH<sub>2</sub> hydrogen bonding. In both of these cases, the geometry about the metal cation in the calculated global minima is a distorted tetrahedral arrangement. Again, this accords well with expectations based on the observed [COOH] binding motifs of the respective monomers.<sup>7</sup> Note, however, that the monomeric Cys•Rb<sup>+</sup> IRMPD spectrum also exhibits peaks that are consistent with the [N,CO,S] binding motif.



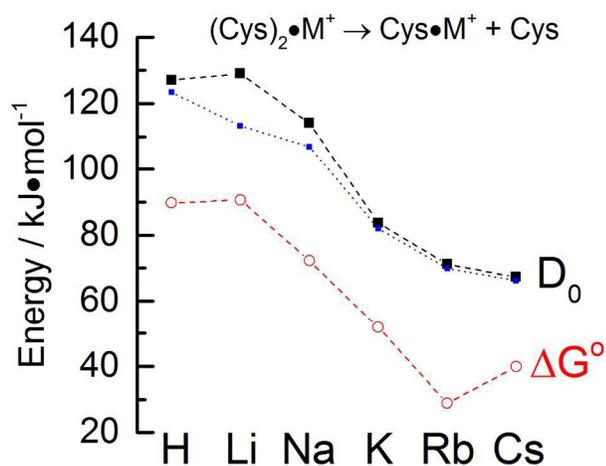
**Figure 7.** The (A) experimental IRMPD spectrum of (Cys)<sub>2</sub>•Na<sup>+</sup> and the calculated harmonic spectra of the (B) global minimum, (C) isomer 4 (4.4 kJ•mol<sup>-1</sup>), (D) isomer 5 (4.4 kJ•mol<sup>-1</sup>), and (E) isomer 9 (12.5 kJ•mol<sup>-1</sup>) structures. The overlaid red trace shows the calculated anharmonic spectrum for the global minimum. Spectra are broadened with a 4 cm<sup>-1</sup> Gaussian linewidth. Calculations were conducted at the B3LYP/6-311++G(d,p) level of theory.

By extending our computational study to include the (Cys)<sub>2</sub>•M<sup>+</sup> (M = K, Rb or Cs) species, we are also able to compare calculated

dissociation energies across the alkali-bound Cys dimer series. This comparison is shown in Figure 9. In all cases, the lowest energy dissociation threshold for the  $(\text{Cys})_2 \cdot \text{M}^+$  species is loss of a neutral Cys. For  $(\text{Cys})_2 \cdot \text{M}^+$  ( $\text{M} = \text{H}, \text{Li}$  or  $\text{Na}$ ), we find  $\Delta G^\circ_{\text{complexation}} \approx 80 \pm 10 \text{ kJ} \cdot \text{mol}^{-1}$ , which implies that 4–7 photons are required for IRMPD in the  $1000 - 2000 \text{ cm}^{-1}$  region. For the heavier members of the alkali-bound Cys dimer series, we find  $\Delta G^\circ_{\text{complexation}} \approx 40 \pm 10 \text{ kJ} \cdot \text{mol}^{-1}$ . This accords well with the difficulties that we experienced in producing and trapping  $(\text{Cys})_2 \cdot \text{M}^+$  ( $\text{M} = \text{K}, \text{Rb}$  or  $\text{Cs}$ ), and it suggests that a more gentle method of production is likely necessary. Note that the  $\text{Rb}^+$ - and  $\text{Cs}^+$ -bound dimer calculations employed a LANL2DZ basis set for the central alkali cation – this will introduce a small amount of error for comparisons with the lighter members of the series. Nevertheless, we are confident that our calculations have qualitatively captured the binding energy trend across the series.



**Figure 8.** The global minima of  $(\text{Cys})_2 \cdot \text{M}^+$  ( $\text{M} = \text{K}, \text{Rb}$  or  $\text{Cs}$ ) as calculated at the B3LYP/6-311++G(d,p) level of theory. The lanl2DZ basis set was used for Rb and Cs. Electronic energies are zero-point corrected.



**Figure 9.** Binding energies for  $(\text{Cys})_2 \cdot \text{M}^+$  ( $\text{M} = \text{H}, \text{Li}, \text{Na}, \text{K}, \text{Rb}$  or  $\text{Cs}$ ) as calculated at the B3LYP/6-311++G(d,p) level of theory. The LANL2DZ basis set was used for Rb and Cs. The blue dotted line shows the results from counterpoise correction for basis set superposition error.

## Conclusions

The proton-, lithium-, and sodium-bound dimers of Cys have been observed for the first time and were studied using IRMPD. By supporting experimental observations with an exhaustive computational study, the structure of  $(\text{Cys})_2 \cdot \text{H}^+$  has been unambiguously identified as an amine-protonated Cys moiety which undergoes intermolecular hydrogen bonding with the carbonyl group of a neutral Cys moiety. The observed spectrum clearly indicates that protonation is localized on an amine group.

In the cases of  $(\text{Cys})_2 \cdot \text{Li}^+$  and (in particular)  $(\text{Cys})_2 \cdot \text{Na}^+$ , the IRMPD spectra suggest the presence of multiple, closely related isomers within the probed ensemble. Comparison with calculated vibrational spectra for the lithium- and sodium-bound species suggest that the dimers adopt structures which consist of the  $[\text{N}, \text{CO}, \text{S}]$  monomeric complexes (first identified by Citir et al.<sup>7</sup>) and a second Cys moiety that interacts with the central alkali cation via the carbonyl oxygen.

Calculations show that the heavier  $(\text{Cys})_2 \cdot \text{M}^+$  ( $\text{M} = \text{K}, \text{Rb}$  or  $\text{Cs}$ ) species are significantly more weakly bound than the proton-, lithium-, and sodium-bound Cys dimers. This accords well with the fact that the heavier alkali-bound Cys dimers could not be produced by our ESI source. In comparing the computed structures for  $(\text{Cys})_2 \cdot \text{M}^+$  ( $\text{M} = \text{K}, \text{Rb}, \text{Cs}$ ) with those of the lighter alkali derivatives, we find that the binding motifs change as the central alkali cation increases in mass (size).  $(\text{Cys})_2 \cdot \text{K}^+$  is predicted to adopt a tridentate/bidentate binding motif (motif M4) for the two Cys moieties, while  $(\text{Cys})_2 \cdot \text{Rb}^+$  and  $(\text{Cys})_2 \cdot \text{Cs}^+$  are both expected to adopt a bidentate/bidentate motif (motif M5). These calculated structures exhibit ligand-metal binding that is similar to that observed for the monomeric  $\text{Cys} \cdot \text{M}^+$  complexes.

## Acknowledgements

We gratefully acknowledge high performance computing support from the SHARCNET consortium of Compute Canada. We are also grateful to the Centre Laser Infrarouge d'Orsay (CLIO) team and technical support staff for the valuable assistance and hospitality. The authors would like to acknowledge financial support from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

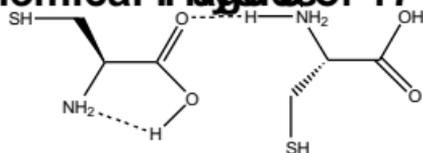
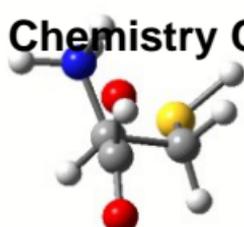
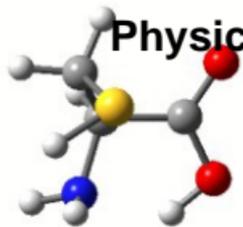
## Notes and references

\*Scott Hopkins: shopkins@uwaterloo.ca

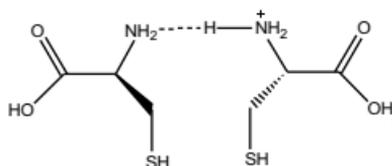
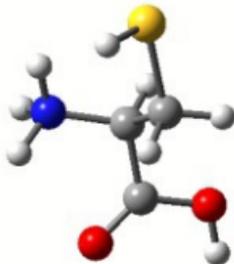
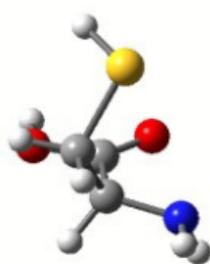
1. Department of Chemistry, University of Waterloo, Waterloo, ON, Canada, N2L 3G1.
  2. Laboratoire Chimie Physique – CLIO, Bâtiment 201, Porte 2, Campus Universitaire d'Orsay, Orsay, France, 91405
- Electronic Supplementary Information (ESI) available:  $(\text{Cys})_2 \cdot \text{M}^+$  ( $\text{M} = \text{H}, \text{Li}, \text{Na}, \text{K}, \text{Rb}$  or  $\text{Cs}$ ) structures are available free of charge. See DOI: 10.1039/b000000x/

1. J. M. Berg and Y. G. Shi, *Science*, 1996, 271, 1081-1085.
2. A. D. Frankel, J. M. Berg and C. O. Pabo, *Proceedings of the National Academy of Sciences of the United States of America*, 1987, 84, 4841-4845.
3. N. M. Giles, A. B. Watts, G. I. Giles, F. H. Fry, J. A. Littlechild and C. Jacob, *Chemistry & Biology*, 2003, 10, 677-693.
4. N. A. Mautjana, J. Estes, J. R. Eyler and A. Brajter-Toth, *Electroanalysis*, 2008, 20, 1959-1967.

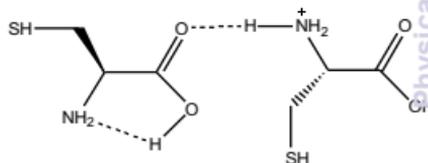
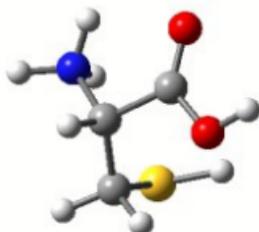
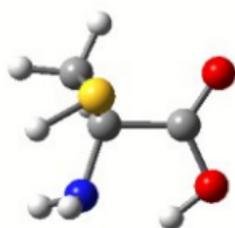
5. J. D. Steill, J. Szczepanski, J. Oomens, J. R. Eyler and A. Brajter-Toth, *Analytical and Bioanalytical Chemistry*, 2011, 399, 2463-2473.
6. T. Meinel, S. Blanquet and F. Dardel, *Journal of Molecular Biology*, 1996, 262, 375-386.
7. M. Citir, E. M. S. Stennett, J. Oomens, J. D. Steill, M. T. Rodgers and P. B. Armentrout, *Int. J. Mass Spectrom.*, 2010, 297, 9-17.
8. 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.
9. P. B. Armentrout, E. I. Armentrout, A. A. Clark, T. E. Cooper, E. M. S. Stennett and D. R. Carl, *Journal of Physical Chemistry B*, 2010, 114, 3927-3937.
10. L. Mac Aleese, A. Simon, T. B. McMahon, J. M. Ortega, D. Scuderi, J. Lemaire and P. Maitre, *Int. J. Mass Spectrom.*, 2006, 249, 14-20.
11. J. M. Ortega, F. Glotin and R. Prazeres, *Infrared Phys. Technol.*, 2006, 49, 133-138.
12. W. S. Hopkins, R. A. Marta and T. B. McMahon, *Journal of Physical Chemistry A*, 2013, 117, 10714-10718.
13. J. K. Martens, I. Compagnon, E. Nicol, T. B. McMahon, C. Clavaguera and G. Ohanessian, *Journal of Physical Chemistry Letters*, 2012, 3, 3320-3324.
14. R. Wu, R. A. Marta, J. K. Martens, K. R. Eldridge and T. B. McMahon, *Journal of the American Society for Mass Spectrometry*, 2011, 22, 1651-1659.
15. 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 and e. al., in *Gaussian, Inc. Wallingford CT 2009*.
16. M. J. Lecours, W. C. T. Chow and W. S. Hopkins, *Journal of Physical Chemistry A*, 2014, 118, 4278-4287.
17. D. J. Wales and J. P. K. Doye, *Journal of Physical Chemistry A*, 1997, 101, 5111-5116.
18. W. S. Hopkins, R. A. Marta, V. Steinmetz and T. B. McMahon, *Phys. Chem. Chem. Phys.*, 2015, 17, 28548-28555.
19. A. D. Becke, *Phys. Rev. A*, 1988, 38, 3098-3100.
20. D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. Debolt, D. Ferguson, G. Seibel and P. Kollman, *Computer Physics Communications*, 1995, 91, 1-41.
21. K. B. Wiberg and P. R. Rablen, *Journal of Computational Chemistry*, 1993, 14, 1504-1518.
22. A. B. McCoy, J. L. Fry, J. S. Francisco, A. K. Mollner and M. Okumura, *Journal of Chemical Physics*, 2005, 122, 14.
23. J. Oomens, J. D. Steill and B. Redlich, *Journal of the American Chemical Society*, 2009, 131, 4310-4319.
24. M. P. Andersson and P. Uvdal, *Journal of Physical Chemistry A*, 2005, 109, 2937-2941.
25. K. T. Crampton, A. I. Rathur, Y. W. Nei, G. Berden, J. Oomens and M. T. Rodgers, *Journal of the American Society for Mass Spectrometry*, 2012, 23, 1469-1478.
26. R. C. Dunbar, J. D. Steill and J. Oomens, *Journal of the American Chemical Society*, 2011, 133, 1212-1215.
27. S. M. Martens, R. A. Marta, J. K. Martens and T. B. McMahon, *Journal of Physical Chemistry A*, 2011, 115, 9837-9844.
28. R. J. Nieckarz, J. Oomens, G. Berden, P. Sagulenko and R. Zenobi, *Phys. Chem. Chem. Phys.*, 2013, 15, 5049-5056.
29. P. Parneix, M. Basire and F. Calvo, *Journal of Physical Chemistry A*, 2013, 117, 3954-3959.



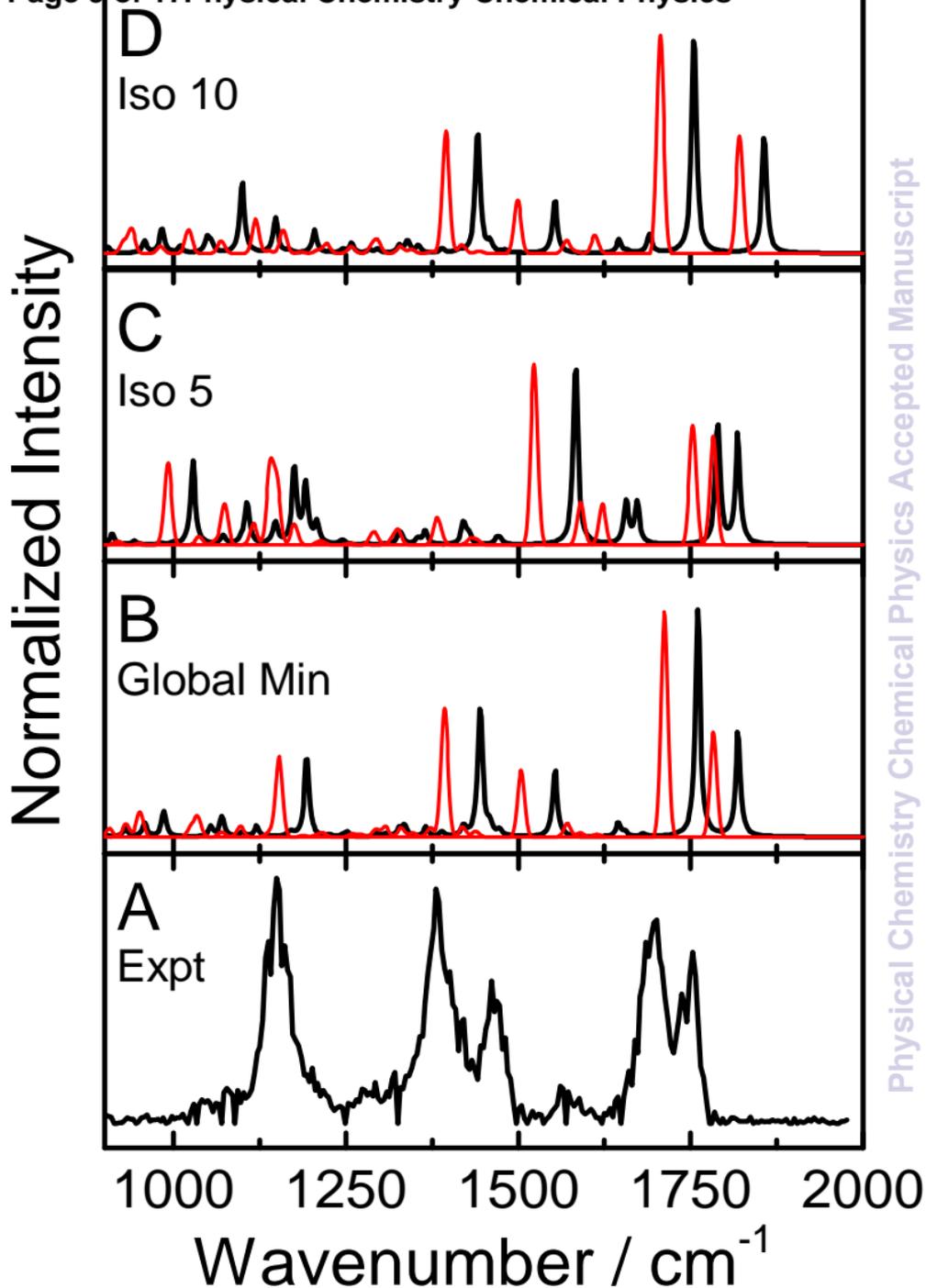
Isomer 10  
(Motif H3;  $17.2 \text{ kJ}\cdot\text{mol}^{-1}$ )

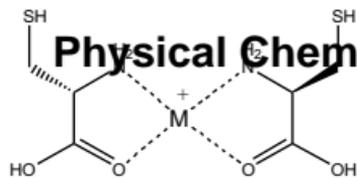


Isomer 5  
(Motif H2;  $11.4 \text{ kJ}\cdot\text{mol}^{-1}$ )

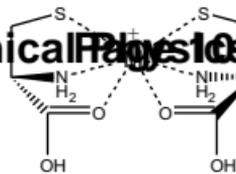


Global Minimum  
(Motif H1;  $0.0 \text{ kJ}\cdot\text{mol}^{-1}$ )

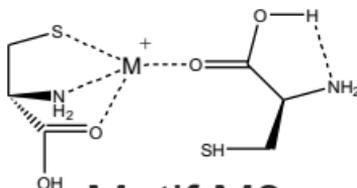




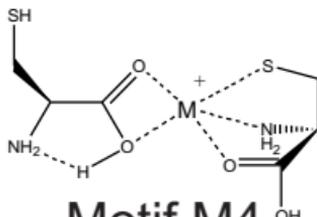
Motif M1



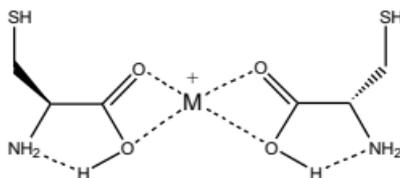
Motif M2



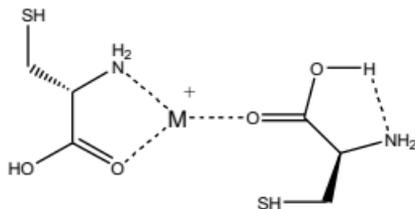
Motif M3



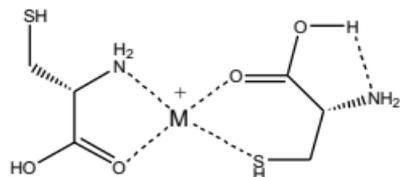
Motif M4



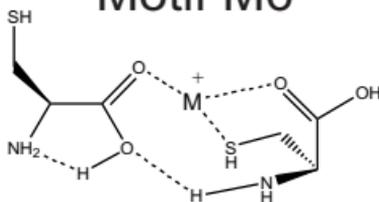
Motif M5



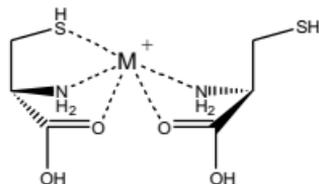
Motif M6



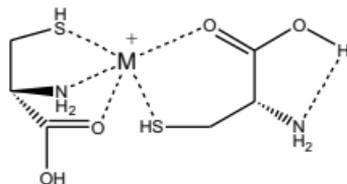
Motif M7



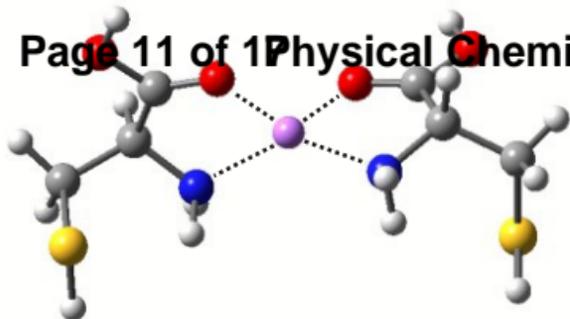
Motif M8



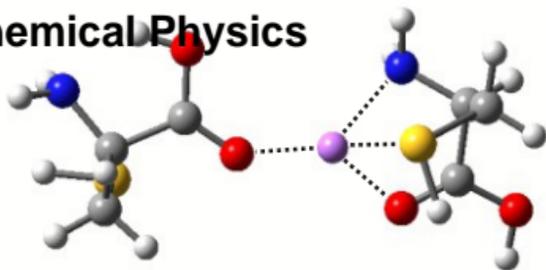
Motif M9



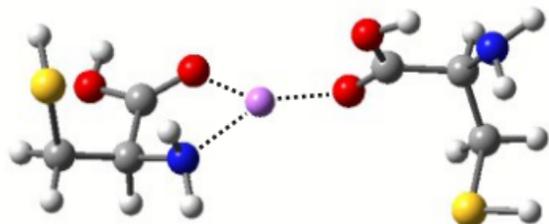
Motif M10



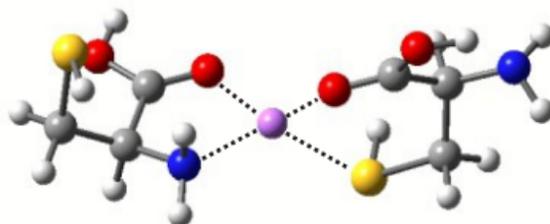
Global Minimum  
(Motif M1; 0.0 kJ·mol<sup>-1</sup>)



Isomer 3  
(Motif M3; 2.8 kJ·mol<sup>-1</sup>)

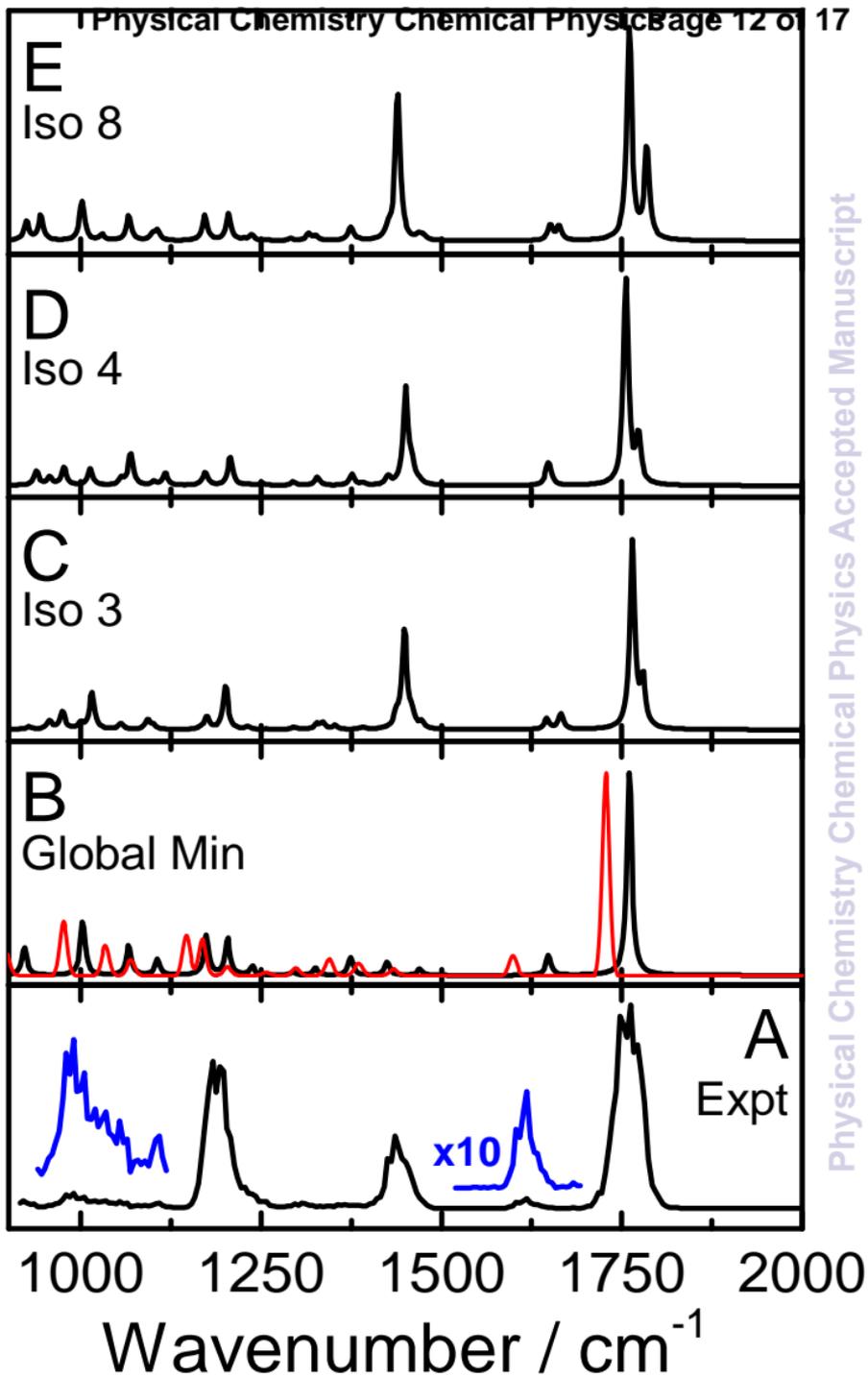


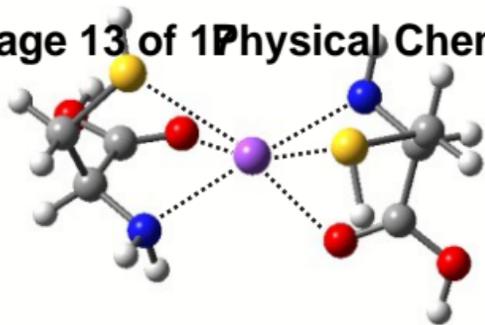
Isomer 4  
(Motif M6; 5.3 kJ·mol<sup>-1</sup>)



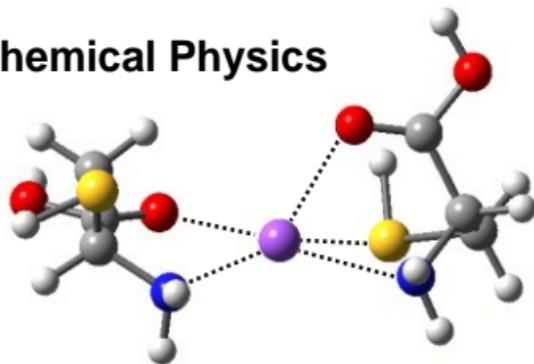
Isomer 8  
(Motif M7; 10.0 kJ·mol<sup>-1</sup>)

Normalized Intensity

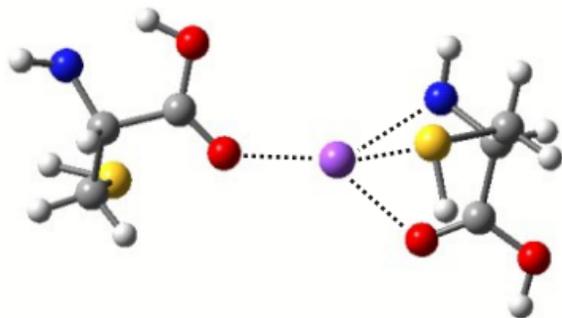




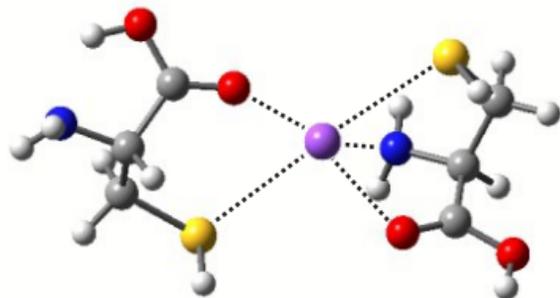
Global Minimum  
(Motif M2;  $0.0 \text{ kJ}\cdot\text{mol}^{-1}$ )



Isomer 4  
(Motif M8;  $4.4 \text{ kJ}\cdot\text{mol}^{-1}$ )

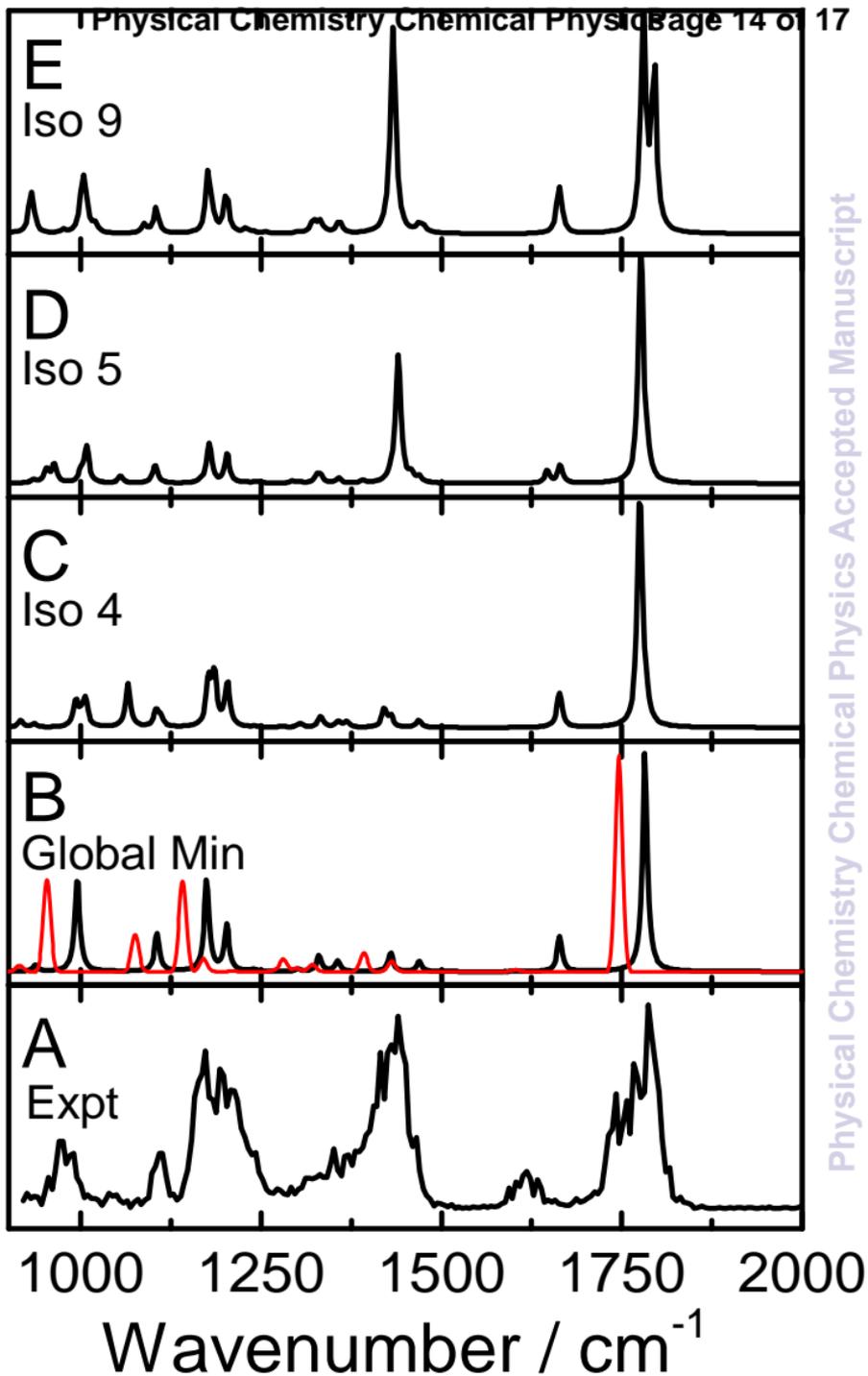


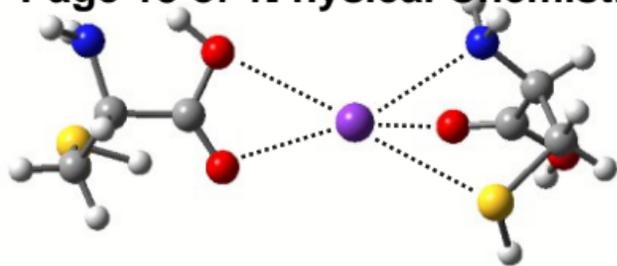
Isomer 5  
(Motif M3;  $4.4 \text{ kJ}\cdot\text{mol}^{-1}$ )



Isomer 9  
(Motif M6;  $12.5 \text{ kJ}\cdot\text{mol}^{-1}$ )

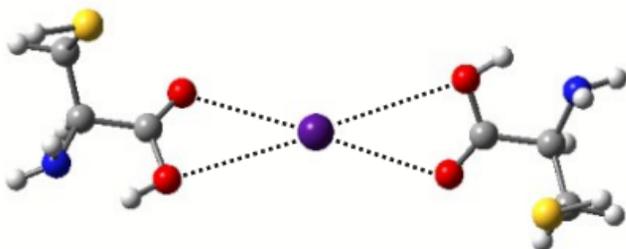
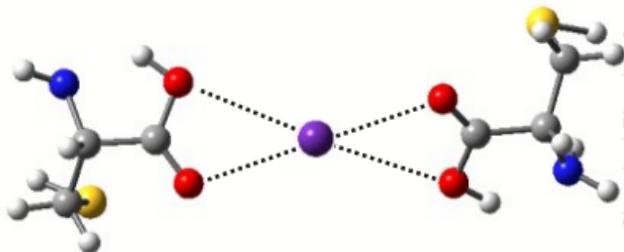
Normalized Intensity



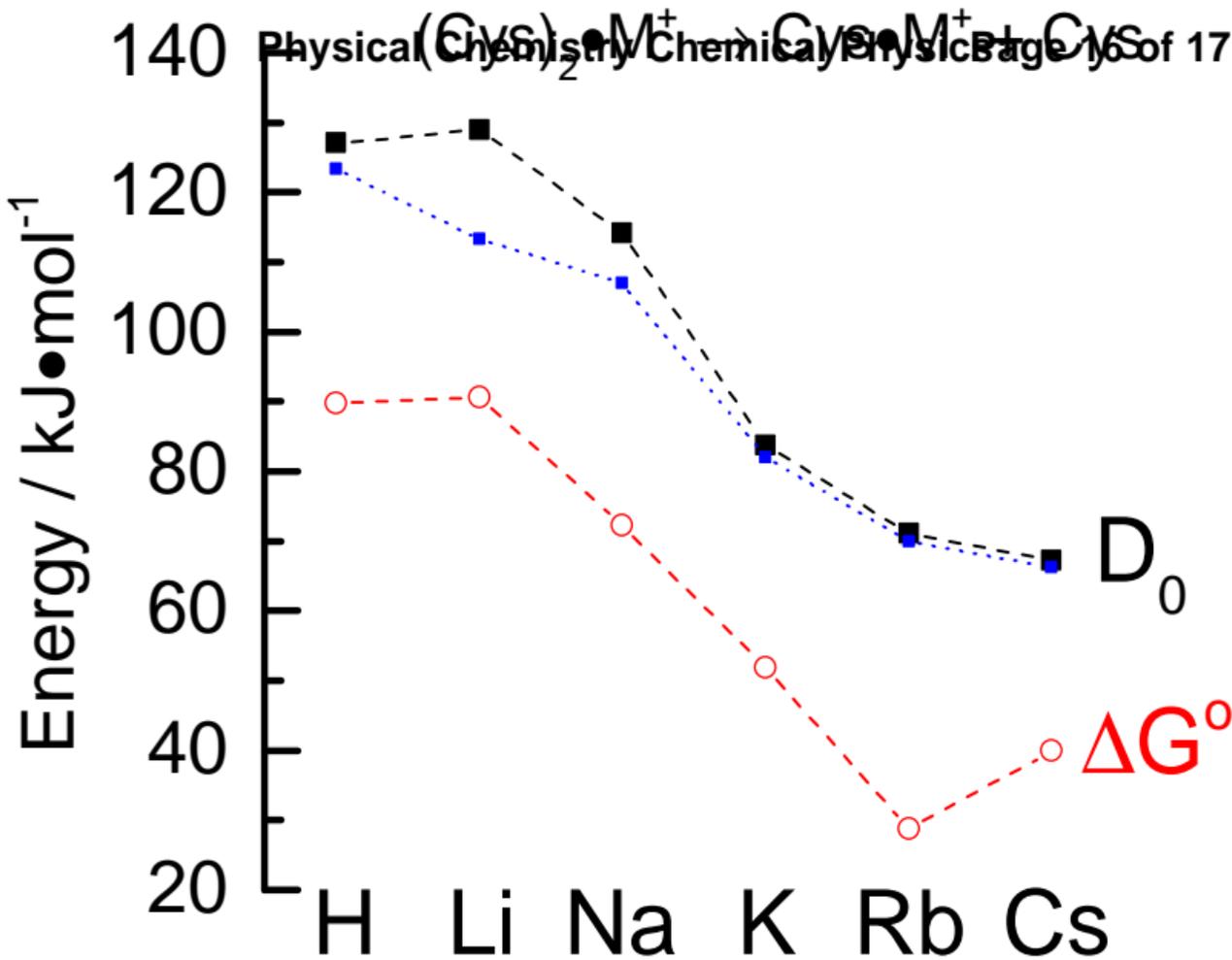


◀  $(\text{Cys})_2 \cdot \text{K}^+$  Global Minimum  
(Motif M4)

$(\text{Cys})_2 \cdot \text{Rb}^+$  Global Minimum ▶  
(Motif M5)



◀  $(\text{Cys})_2 \cdot \text{Cs}^+$  Global Minimum  
(Motif M5)



# Physical Chemistry Chemical Physics

