Gas-phase Structures Reflect Pain-relief Potency of Enkephalin Peptides.

Aleksandr Y. Pereverzev,[1] István Szabó,[2] Vladimir N. Kopysov[1], Edina Rosta[2], and Oleg V. Boyarkin[1*]

[1]Laboratoire de Chimie Physique Moléculaire, École Polytechnique Fédérale de Lausanne, Station-6, 1015 Lausanne, Switzerland

[2]Department of Chemistry, King's College London, Britannia House, 7 Trinity Street, London SE1 1DB, United Kingdom.

Abstract

The geometry of drug molecules at their targeted binding sites *in vivo* is a key for predicting their pharmacological potency but often challenging to determine. Instead, some empirical rules may relate drug efficiency to certain "pharmacological" parameters of molecular structures in a condense phase. These rules fail however for pain-relief peptide drugs enkeplalins, which bind into hydrophobic pockets of opioid receptors and exhibit a variety of conformational states in condense phase. We use cold ion spectroscopy and high-level quantum-chemical computations to solve structures of conformers of some enkephalin peptides in the gas phase. The derived gas-phase structural parameters clearly correlate with the known pharmacological efficiency of the studied drugs. This result allows us to propose a rational explanation for the empirical rules and suggests that high structural resolution of gas-phase methods, perhaps, can be used for predicting the relative potency of ligand drugs that target hydrophobic pockets of receptors.

* Corresponding Author: oleg.boiarkin@epfl.ch

The recent advances in the gas-phase cold ion spectroscopy and computational methods of quantum chemistry make it now possible to solve the accurate intrinsic structures of small to midsize protonated biomolecules, including peptides, carbohydrates, drug molecules, etc.[1, 2] The relation between these dehydrated structures and the native structures that biomolecules adopt in solution often remains unclear. This circumstance challenges a use of gas phase spectroscopy for solving biologically relevant problems, in particular, for *in silico* drug design. Studying microsolvated biomolecules is likely one possible compromised solution of the problem.[3] Another case where intrinsic structures might be suitable for biology is when the native environment of drugs at their active sites is not a solution. In this report we explore the relevance of gas-phase structures to biological activity for those small peptide ligands that in vivo bind into hydrophobic environment. The pharmacological activity of many drugs is crucially determined by the geometry they adopt *in vivo* upon binding to targeted proteins. Over decades, researchers developed some semi-empirical rules that correlate pharmacological activity of drugs with their structures in condensed phase measured by NMR[4] or by X-ray crystallography.[5] Enkephalins are known pain-relief drugs that bind into hydrophobic pockets of transmembrane opioid protein receptors.[6-8] These pentapeptides differ by some residues and/or by their stereoisomeric form (*L*/*D*), but all contain two aromatic residues, Phe and Tyr. The empirical rules suggest that an efficient opioid drug has to have short distances between the two aromatic rings and between the rings and the protonated amine.[5, 9-11] The pharmacological activity of Ala^2-enkephalin, YAGFL, with all the residues in *L* stereoisomeric form (denoted here as *LL)* increases substantially upon the replacement of *L*-Ala^2 by its *D*-enantiomer (*DL*) and becomes even higher for the [*D*-Ala^2-*D*-Leu^5]-Enk stereoisomer (*DD*), also known as *DADLE*.[12, 13] Studies in condensed phase revealed no correlation between the activity of the three YAGFL stereoisomers and their interchromophore distances.[9] Recent studies of enkephalins using gas-phase spectroscopy and quantum chemistry calculations determined that the interchromophore spacing is, indeed, large (11.6 Å) in low potent opioid peptides YGGFL (denoted here as *Enk*) and YAGFL.[14, 15] Consistently with the pharmacological rules, this distance estimated from the gas-phase FRET measurements for one of the conformers of *DD* is short (< 7 Å).[16] Although the ligand environments in a hydrophobic pocket and in vacuum differ by weak hydrophobic interactions, the much stronger hydrophilic couplings with the environment are limited in the first and absent in the second case. This might make the gas phase rather than condensed phase more appropriate for mimicking the environment of the pockets and predicting the drug efficacy. In order to explore this hypothesis, herein we explicitly solve the intrinsic structures of two protonated stereoisomeric YAGFL

peptides, *DD* and *DL*, using a combination of conformer-selective cold ion IR/UV spectroscopy,[2] high-level quantum-chemical calculations, and an extensive conformational search. We then compare the pharmacologically critical distances in these and in *LL* and *Enk* opioid peptides with their pharmacological potency and propose a model that rationalizes the old empirical pharmacological rules.

Our tandem mass-spectrometer[17-19] (see SI for details) enables measuring five types of photofragmentation spectra: electronic UV and IR-*UV* hole burning,[20] vibrational *IR*-UV gain,[19, 21] *IR*-UV depletion[17, 22, 23] and IR-*IR*-UV hole burning[18, 20] (here and below *italic* font in the names of multi-laser techniques designates the scanned laser). The derived conformer-specific IR spectra were compared with the vibrational spectra computed for pre-selected low-energy calculated geometries of a molecule. The reliability of the workflow of conformational search and the accuracy of calculations has been confirmed by reproducing the most stable structure and IR spectrum of *LL* isomer[14] (see SI for details).

Figure 1a shows the all-conformer *IR*-UV gain spectrum of *DD*. The large number of the observed transitions indicates the presence of, at least, three conformers that are highly abundant in the gas phase under the conditions of our experiment. This is further confirmed by the IR-*UV* hole-burning spectra of *DD* (Fig. S6). We name these conformers as *DD*-**A**, *DD*-**B**, and *DD*-**C**, respectively. The electronic spectrum of *DD* (Fig. S6) appears highly congested, prohibiting a use of *IR*-UV depletion technique for this stereoisomer. Instead, the more complex IR-*IR*-UV hole-burning technique was employed for decomposing the all-conformer gain spectrum into the separate, conformer-specific ones.

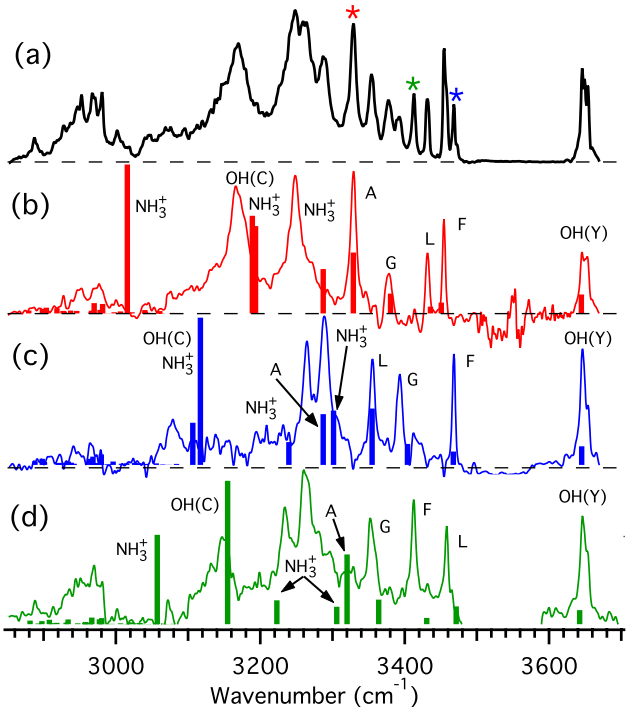

Figure 1. a) *IR*-UV gain spectrum of cold *DD*. Colored asterisks denote transitions used for IR-*IR*-UV hole-burning; b)-d) conformer-selective spectra of *DD* obtained by IR-*IR*-UV hole-burning spectroscopy. Vertical lines show the transitions in the best-matching spectra calculated for the computed *DD* structures (shown in Figure 2). The UV laser in all the experiments was fixed at 35370 cm^{-1}. Calculated NH-stretch transitions of residues and of N-terminus are labeled by their symbols.

Figures 1b-d show the IR-*IR*-UV spectra measured with pump IR OPO fixed at three different IR transitions of *DD*. All strong transitions in the gain spectrum of *DD* (Fig. 1a) do appear in one of these conformer-specific spectra. This suggests that all highly abundant conformers present in the experiment were found. Comparison of the IR spectra of these three conformers, named *DD*-**A**, *DD*-**B**, and *DD*-**C**, with the IR transitions computed for a pool of pre-selected candidate geometries (Figure S2) validates the three structures, which exhibit the best "exp-calc" spectral match (Figures 1b-d). All of them (Figures 2a-c) adopt the compact double bend motive[9] with all "pharmacological" distances being short (Table S3). The earlier FRET measurements also predicted a small spacing between the two chromophore rings in this isomer, although only one conformer could be assessed by this method.[16]

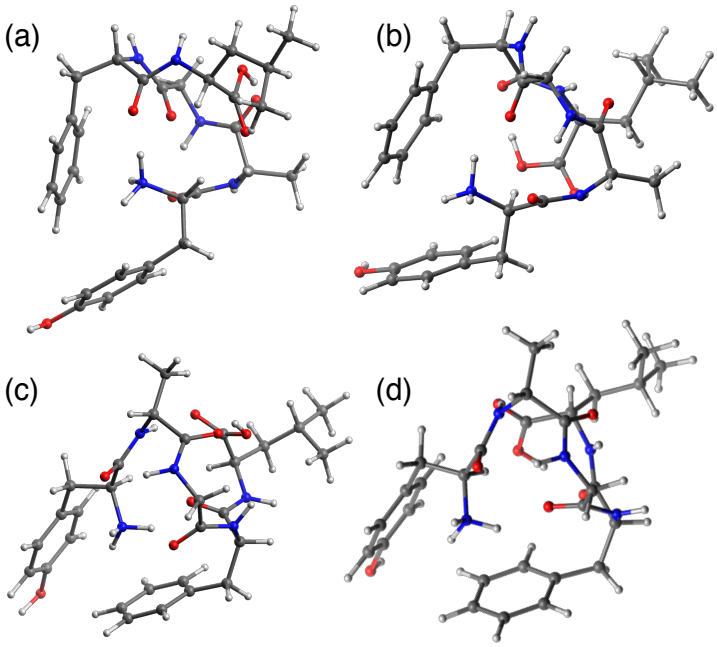

Figure 2. Structures of the conformers (a) *DD*-**A**, (b) *DD*-**B**, (c) *DD*-**C** and (d) *DL*-**B** for which the calculated vibrational spectra exhibit the best match to the experimental ones; the calculated relative potential energies of the *DD* structures are 0.0, 0.23 and 5.1 kcal/mol, respectively.

The third stereoisomer is the most intriguing, since the pharmacological potency of *DL* is higher than of *Enk* and *LL*, but lower than that of *DD*. The FRET measurements failed to estimate the interchromophore distance in *DL* within the required certainty (Fig. S7).

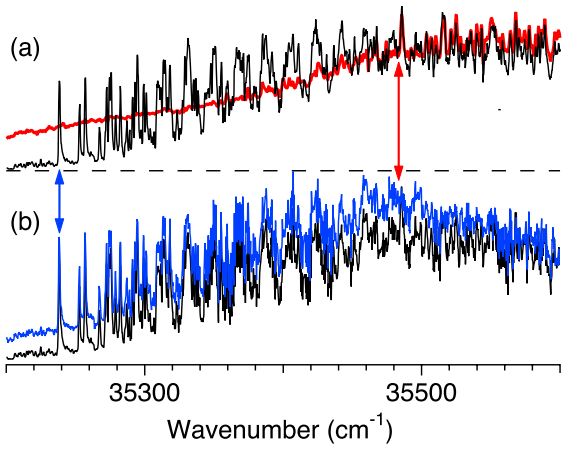

Figure 3. Electronic spectrum of *DL* (black traces), and its IR-*UV* hole-burning spectra recorded with IR wavenumber fixed at (a) 3411.2 and (b) 3466.5 cm^{-1} (red and blue traces, respectively). The arrows marks the electronic band origin in *DL*-**A** (blue) and *DL*-**B** (red).

The UV spectrum of this isomer (Fig. 3) resembles by its complexity a combination of the spectra of *LL* and *DD*.[16] The number of peaks in the *IR*-UV gain spectrum of *DL* (Fig. 4a) suggests the contributions from, at least, two abundant conformers. IR-*UV* hole-burning spectra (Fig. 3) confirm the presence of only two highly abundant conformers (conformational

families) of *DL* with their electronic band origins at 35238 cm^{-1} (*DL*-**A**) and 35486 cm^{-1} (*DL*-**B**). The sharp peaks removed by IR pre-heating of *DL*-**A** (Fig. 3a), and the remaining peaks, suppressed by IR pre-excitation of *DL*-**B**, have been assigned to each of these conformers. Note that, due to the large separation of their band origins, the relative intensities of similar IR transitions in the two conformers in the gain spectrum of *DL* (Fig. 4a) do not correctly reflect their relative abundance, such that the abundance of *DL*-**B** can be higher than that of *DL*-**A**. The observed vibrational resolution in the UV spectrum of *DL*-**A** enabled conformer-selective *IR*-UV depletion spectroscopy of this conformer (Fig. 4b), while IR-*IR*-UV technique was required to attain conformational selectivity with *DL*-**B** (Fig. 4c), whose UV spectrum is too congested.

Among the spectra calculated for all 20 computed preselected candidate structures of *DL* that have relative energies below 6 kcal/mol (Fig. S4), the IR spectrum of the lowest energy structure exhibits the best match to the measured spectrum of *DL*-**B** (Fig. 4c). The match is quite good with a typical for a harmonic approximation of theory overestimate of frequencies for the strongly red-shifted NH-stretch transitions. We thus assign the *DL*-**B** conformer to the most stable calculated structure of *DL* (Fig. 2d). In many aspects the geometry of *DL*-**B** is similar to the three validated structures of *DD*. In all of them the N-terminus forms proton-π bonds with two aromatic rings and a hydrogen bond with amide group of the Ala residue. Similarly to the structures of *DD*, this characteristic arrangement makes the *DL*-**B** structure also "compact" with the interchromophore spacing of 4.7 Å. Spectroscopically, the NH_3^+ bindings in the four structures reveal themselves as only moderate redshifts for the two of its NH-stretch transitions, but a large redshift for the third one (H-bond).

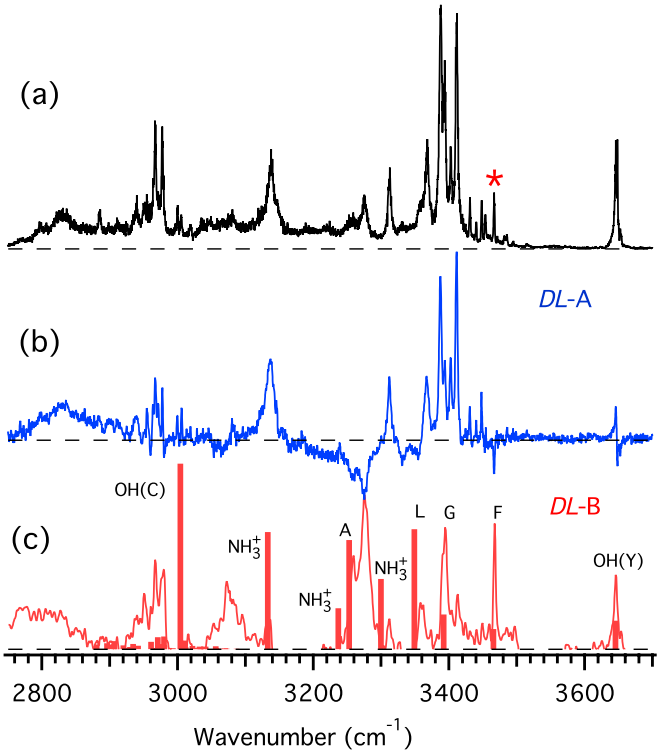

Figure 4. (a) *IR*-UV gain, (b) *IR*-UV depletion and (c) IR-*IR*-UV hole-burning spectra of *DL* isomer of protonated peptide YAGFL. UV wavenumber was fixed at 35140 cm^{-1} in (a) and (c), and at 35238 cm^{-1} in (b). Pump IR OPO in (c) was fixed at 3466.5 cm^{-1} (red asterisk in (a)) Red sticks in (c) indicate the calculated scaled vibrational transitions in the lowest-energy computed structure *DL*-**B** (Fig. 2d).

Overall, the calculated lowest-energy structures of all three stereoisomers correspond to the experimentally observed conformers, giving us full confidence in high reliability of our calculations. Nevertheless, none of the spectra calculated for the 20 candidate structures of *DL* with the energies below 6 kcal/mol exhibit a satisfactory match to the spectrum of *DL*-**A** conformer (Fig. S4). Further calculations of IR spectra for selected 27 more structures that represent all the structural clusters with the energies below 12 kcal/mol still did not reveal a satisfactory match to the experiment (Fig. S8). We may suggest that this conformer is not among the lowest gas-phase structures but is rather a kinetically trapped solution-like geometry.[24-27] Under the typical "soft" conditions in our electrospray source (optimal low amplitude of RF field in the ion funnel), which are used to minimize in-source fragmentation, *DL*-**A** could not fully overcome the energy barrier to relax into an intrinsic structure after desolvation in the source and before collisional thermalization in the subsequent (room temperature) accumulation ion trap. Increasing RF field in the source heats up the ions and results in a noticeable reduction of relative intensities of the IR transitions due to this conformer

in the gain spectrum (Figure S9). This observation implies that *DL*-**A** is indeed trapped in a local potential energy well and has to overcome a barrier to interconvert to the most stable gas phase structure *DL*-**B**.

Both the UV and IR spectra suggest that *DL*-**A** should differ substantially from the *DD* and *DL*-**B** structures and is, likely, non-compact. Similar to *LL*, in which NH-stretches of three residues are, essentially, free of H-bonding[14] and appear within a narrow interval of 30 cm^{-1} above ~3377 cm^{-1} (Fig. S5), the NH-stretches of three amides in *DL*-**A** are concentrated within just 25 cm^{-1} above ~3387 cm^{-1} (Fig. 4). Conversely, in *DD* and *DL*-**B** structures the three high frequency NH stretches are spread over 75 to 112 cm^{-1} intervals, reflecting the diverse strength of H-bonds they are involved in. This comparison is consistent with a non-compact rather than a compact structure of *DL*-**A**. Also, among the all studied stereoisomers of $[YAGFL+H]^+$, the electronic band origin in *DL*-**A** is the most red-shifted to almost that in $[YA+H]^+$.[17, 28] In this dipeptide the protonated N-terminus is strongly coupled to the aromatic ring, but exhibits no other significant non-covalent interactions. This implies a similar charge solvation by the Tyr ring in *DL*-**A** and hence suggests that this conformer was not able to assemble into a compact structure around the N-terminus.

Overall, the low potent drugs, *LL* and YGGFL, exhibit single intrinsic structures, characterized by a large, 11.6 Å, interchromophore spacing. The three detected gas-phase conformers of the most potent drug, *DD,* appear all compact with this distance being as short as 4.4-4.7 Å. The mid-efficient drug, *DL*, resides in the gas phase in two conformational states, the energetically lowest of which is also compact and resembles the conformers of *DD*. This conformer thus should exhibit a pharmacology that is similar to *DD*. The second *DL* conformer is "kinetically trapped" in a solution-based state, which is most likely non-compact and therefore pharmacologically less efficient. With this the only uncertainty, the empirical rules of "pharmacological" distances, when applied to gas-phase structures, properly reflect the pharmacological potency of the studied enkephalins. A reasonable question remains why these rules applied to condense phase opiates, which are zwitterions at physiological pH, and to the desolvated protonated peptides should work at all? Indeed, at their binding sites the structures of opiates (including enkephalins), which were solved in benchmark experiments[8, 29] and by computer docking of the ligands to opioid receptors,[30] differ significantly from that in condense and the gas phase. Energetically competing with desolvation, even the ionic state of the drugs in the receptor binding pockets can change (e.g from neutral zwitterion to protonated).[31] What however appears essential and common to all the opiates is the pharmacophore, which in enkephalins consist of a Tyr side chain and a nearby NH_3^+.[29] It is exactly this group that buries

deeply into the binding pockets and determines in first the pharmacological efficacy of the ligands. Compared with cleavage of other types of H-bonds, desolvation of proton is energetically very costly. This prompts us to propose that the geometries of enkephalins, in which the proton is protected from hydration, exhibit lower energy barrier to adopt the final structure at the receptors. The determined herein geometries of *DD* and *DL*-**B**, in which the proton is surrounded by two large hydrophobic aromatic rings comply with this constraint. The pharmacological rules empirically rationalized such structural motive of pharmacophores for condense phase: short interchromophore and chromophore–proton distances imply (although non-exclusively) a protonated N-terminus protected by aromatics ring from solvation. Solution phase conformers with "protected" proton are most likely to retain this motive upon desolvation in an electrospray ion source, while the structures with a solvated proton will have to undergo significant conformational changes. In this respect, gas-phase structures properly reflect the structural motive of "protected" pharmacophores in solutions. The crucial role of dehydration in energetics of ligand–receptor bindings is well recognized, but cannot yet be fully accounted for.[32, 33] A detection of compact conformations in the gas phase may serve as a probe of opioid drug efficacy – a subject of further verifications with different types of opioids. The continuously improving accuracy and reliability of quantum chemistry calculations (yet validated by spectroscopy) can, potentially, make solving intrinsic structures a good alternative to computer docking for evaluating binding efficiency of ligands into hydrophobic pockets of acceptor proteins.

Acknowledgements

This work was supported by Swiss National Science Foundation (grants 206021_164101, 200020_172522). We thank Dr. M. Audagnotto and prof. M. Dal Peraro for discussions of ligand docking processes.

References

1. T. R. Rizzo and O. V. Boyarkin, *Gas-Phase Ir Spectroscopy and Structure of Biological Molecules*, 2015, **364**, 43-97.
2. O. V. Boyarkin, *International Reviews in Physical Chemistry*, 2018, **37**, 559-606.
3. N. S. Nagornova, T. R. Rizzo and O. V. Boyarkin, *Science*, 2012, **336**, 320-323.
4. N. Collins, J. L. FlippenAnderson, R. C. Haaseth, J. R. Deschamps, C. George, R. Kover and V. J. Hruby, *J. Am. Chem. Soc.*, 1996, **118**, 2143-2152.

5. J. R. Deschamps, J. L. Flippen-Anderson and C. George, *Biopolymers*, 2002, **66**, 287-293.
6. A. Janecka, J. Fichna and T. Janecki, *Curr Top Med Chem*, 2004, **4**, 1-17.
7. M. Waldhoer, S. E. Bartlett and J. L. Whistler, *Annual Review of Biochemistry*, 2004, **73**, 953-990.
8. A. Manglik, A. C. Kruse, T. S. Kobilka, F. S. Thian, J. M. Mathiesen, R. K. Sunahara, L. Pardo, W. I. Weis, B. K. Kobilka and S. Granier, *Nature*, 2012, **485**, 321-U170.
9. J. R. Deschamps, C. George and J. L. FlippenAnderson, *Biopolymers*, 1996, **40**, 121-139.
10. B. G. Nielsen, M. O. Jensen and H. G. Bohr, *Biopolymers*, 2003, **71**, 577-592.
11. R. Schwyzer, *Biochemistry-US*, 1986, **25**, 6335-6342.
12. H. W. Kosterlitz, J. A. H. Lord, S. J. Paterson and A. A. Waterfield, *Brit J Pharmacol*, 1980, **68**, 333-342.
13. J. S. Morley, *Annu Rev Pharmacol*, 1980, **20**, 81-110.
14. T. K. Roy, V. Kopysov, A. Pereverzev, J. Šebek, R. B. Gerber and O. V. Boyarkin, *Phys. Chem. Chem. Phys.*, 2018, **20**, 24894-24901.
15. N. L. Burke, J. G. Redwine, J. C. Dean, S. A. McLuckey and T. S. Zwier, *Int. J. Mass Spectrom.*, 2015, **378**, 196-205.
16. V. Kopysov and O. V. Boyarkin, *Angew. Chem. Int. Ed.*, 2016, **55**, 689-692.
17. J. A. Stearns, M. Guidi, O. V. Boyarkin and T. R. Rizzo, *J. Chem. Phys.*, 2007, **127**, 154322-154327.
18. A. Y. Pereverzev, X. Cheng, N. S. Nagornova, D. L. Reese, R. P. Steele and O. V. Boyarkin, *J. Phys. Chem A*, 2016, **120**, 5598-5608.
19. A. Y. Pereverzev and O. V. Boyarkin, *Phys. Chem. Chem. Phys.*, 2017, **19**, 3468-3472.
20. V. A. Shubert and T. S. Zwier, *J. Phys. Chem. A*, 2007, **111**, 13283-13286.
21. Y. Inokuchi, O. V. Boyarkin, R. Kusaka, T. Haino, T. Ebata and T. R. Rizzo, *J. Am. Chem. Soc.*, 2011, **133**, 12256-12263.
22. L. C. Snoek, R. T. Kroemer and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2002, **4**, 2130-2139.
23. P. M. Bialach, T. C. Martin and M. Gerhards, *Phys. Chem. Chem. Phys.*, 2012, **14**, 8185-8191.
24. J. A. Silveira, K. L. Fort, D. Kim, K. A. Servage, N. A. Pierson, D. E. Clemmer and D. H. Russell, *J. Am. Chem. Soc.*, 2013, **135**, 19147-19153.
25. S. Warnke, J. Seo, J. Boschmans, F. Sobott, J. H. Scrivens, C. Bleiholder, M. T. Bowers, S. Gewinner, W. Schollkopf, K. Pagel and G. von Helden, *J. Am. Chem. Soc.*, 2015, **137**, 4236-4242.
26. A. L. Patrick, A. P. Cismesia, L. F. Tesler and N. C. Polfer, *Int. J. Mass Spectrom.*, 2017, **418**, 148-155.
27. H. Wako, S. Ishiuchi, D. Kato, G. Feraud, C. Dedonder-Lardeux, C. Jouvet and M. Fujii, *Phys. Chem. Chem. Phys.*, 2017, **19**, 10777-10785.
28. V. Kopysov, N. S. Nagomova and O. V. Boyarkin, *J. Am. Chem. Soc.*, 2014, **136**, 9288-9291.
29. S. Granier, A. Manglik, A. C. Kruse, T. S. Kobilka, F. S. Thian, W. I. Weis and B. K. Kobilka, *Nature*, 2012, **485**, 400-U171.
30. F. Wieberneit, A. Korste, H. B. Albada, N. Metzler-Nolte and R. Stoll, *Dalton Trans.*, 2013, **42**, 9799-9802.
31. A. V. Onufriev and E. Alexov, *Q. Rev. Biophys.*, 2013, **46**, 181-209.
32. R. Abel, T. Young, R. Farid, B. J. Berne and R. A. Friesner, *J. Am. Chem. Soc.*, 2008, **130**, 2817-2831.
33. M. E. Wall, G. Calabro, C. I. Bayly, D. L. Mobley and G. L. Warren, *J. Am. Chem. Soc.*, 2019, **141**, 4711-4720.