Gas-phase probes of kinetically trapped peptides

THÈSE N° 7301 (2016)

PRÉSENTÉE LE 11 NOVEMBRE 2016
À LA FACULTÉ DES SCIENCES DE BASE
LABORATOIRE DE CHIMIE PHYSIQUE MOLÉCULAIRE
PROGRAMME DOCTORAL EN CHIMIE ET GÉNIE CHIMIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Liudmila VORONINA

acceptée sur proposition du jury:

Prof. J. Waser, président du jury
Prof. T. Rizzo, directeur de thèse
Prof. V. Wysocki, rapporteuse
Dr G. von Helden, rapporteur
Prof. U. Lorenz, rapporteur

EPFL

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2016

Abstract

Gas-phase techniques are promising for studying intrinsically disordered peptides and proteins, as they can be implemented in a conformer-selective way. One of the issues that prevent them from finding a broader use is the question of to which extent the structure of a biological molecule is preserved when transferring it from solution to the gas phase *via* electrospray. We focus on the proline-rich peptides and *cis/trans* isomerization of prolyl-peptide bonds, which is known to have high barrier on the potential energy surface. Upon electrospraying, solution-like structures can find themselves kinetically trapped behind these barriers, as we explicitly show for the case of bradykinin 1-5 fragment. Using nuclear magnetic resonance, we determine that it adopts mostly all-*trans* conformation in solution, while three distinct conformational families are observed in the ion mobility experiment in the gas phase. Cryogenic ion spectroscopy combined with first-principles simulations allows us to identify the major conformers and show that the solution structural preferences of the prolyl-peptide bonds are preserved in the kinetically trapped conformational family, while the lowest-energy gas-phase structures have one of the prolines in the *cis* conformation. We show how collisional cross-sections and infrared spectra are used to guide and verify the calculations, giving rise to a new type of symbiosis between theory and experiment. This is especially valuable for the kinetically trapped structures, as in this case the energy criterion cannot be decisive.

As a next step, we characterize the full nonapeptide bradykinin in the +3 charge state. Using field-asymmetric ion mobility spectrometry combined with cryogenic ion spectroscopy, we demonstrate the presence of three major conformational families in the gas phase, one of which is kinetically trapped. This result is in agreement with the drift-tube ion mobility data published previously, and we propose a correspondence between the conformational families separated by field-asymmetric and drift-tube ion mobility spectrometry. We obtain conformer-specific infrared spectra of the major conformers and assign the vibrational bands using ^{15}N isotopic labeling. Substituting carbon atoms in the phenyl ring with their ^{13}C isotope allows us to separate two types of structures according to

their fragmentation pattern upon electronic excitation. This method can serve as a complementary way to introduce conformer-selectivity when studying large molecules.

Finally, we assess the result of substituting one or two prolines in the bradykinin sequence with alanine. This method has been used previously to assign the peaks in the drift-time distributions to certain conformations of the prolyl-peptide bonds. We show that the mutants do reproduce a part of the conformers of the original peptide, but also form additional structures due to the higher flexibility of the alanine backbone compared to proline. We conclude that the proline-to-non-proline substitutions are helpful to assign the structures, but have to be used in conjunction with spectroscopic techniques, which allow detailed comparison of the structures of the mutant and the native peptide.

In general, this work is one of the first steps towards a database of atomic-resolution peptide structures in the gas phase and a better understanding of the capabilities and limits of spectroscopy and ion mobility in the gas phase applied to intrinsically disordered peptides.

Zusammenfassung

Bei der Erforschung intrinsisch ungeordneter Peptide und Proteine sind Untersuchungen in der Gasphase vielversprechend, da sie konformeren-selektiv sein können. Einer der Gründe, die bisher eine breite Anwendung solcher Techniken verhindert haben, ist die Problematik möglicher Strukturveränderungen eines biologischen Moleküls beim Transfer aus der Lösung in die Gasphase, z.B. mittels Electrospray. In der vorliegenden Arbeit konzentrieren wir uns auf Prolin-reiche Peptide und auf die *cis/trans* Isomerisierung von Prolyl-Peptid-Bindungen, welche eine hohe Energiebarriere auf der Potentialhyperfläche aufweist. Die in Lösung vorliegenden Strukturen können selbst nach dem Eletrospray-Prozess auch in der Gasphase kinetisch gefangen sein, was hier im Fall des doppelt protonierten Bradykinin 1-5 Fragmentes demonstriert wird. Anhand von Kernspinresonanzspektroskopie bestimmen wir, dass dieses Fragment in der Lösung hauptsächlich die *all-trans* Konformation annimmt, während in der Gasphase mittels Ionen-Mobilitäts-Experimenten drei unterschiedliche Konformationsfamilien beobachtet werden. Kryogenische Ionenspektroskopie in Kombination mit quantenchemischen Simulationen helfen bei der Identifikation der dominierenden Konformere und zeigen, dass die in Lösung beobachteten strukturellen Vorlieben der Prolyl-Peptid-Bindungen auch in der kinetisch gefangenen Konformationsfamilie gesehen werden, während in den energetisch stabilsten Strukturen in der Gasphase eines der Proline in *cis* Konformation vorliegt. Wir zeigen, wie Stossquerschnitte und Infrarotspektren als Leitlinien und Verifizierung von quantenmechanischen Rechnungen verwendet werden können. Dies stellt eine neue Symbiose zwischen Theorie und Experiment dar. Insbesondere im Falle der kinetisch gefangenen Strukturen ist dies von Bedeutung, da die Energie nicht das entscheidende Kriterium ist.

Nach der Untersuchung des Konformationsraums des Bradykinin 1-5 Fragmentes, charakterisieren wir das komplette Nonapeptid Bradykinin im dreifach positiv geladenen Zustand. Mit Feld-asymmetrischer Ionenmobilitätsspektrometrie in Kombination mit kryogenischer Ionenspektroskopie demonstrieren wir die Existenz dreier

Hauptkonformationsfamilien in der Gasphase, wovon eine kinetisch gefangen ist. Dieses Ergebnis stimmt überein mit den bereits publizierten Daten aus Drift-Rohr Ionenmobilitäts-Experimenten. Wir vermuten einen Zusammenhang zwischen den Konformationsfamilien die durch Feld-asymmetrische Ionenmobilitätsspektrometrie beziehungsweise Drift-Rohr Ionenmobilität aufgetrennt wurden. Wir erhalten Konformeren-spezifische Infrarotspektren der Hauptkonformere und ordnen die Bänder anhand von ^{15}N Isotopenmarkierung zu. Indem Kohlenstoffatome im Phenylring mit deren ^{13}C Isotop ersetzt werden, können zwei Strukturtypen anhand ihrer Fragmentierungsmuster nach elektronischer Anregung getrennt werden. Diese Methode kann als weiterer Weg zur Konformeren-Selektivität dienen, insbesondere bei Untersuchungen grosser Moleküle.

Schlussendlich beurteilen wir das Resultat, wenn eines oder zwei der Proline in Bradykinin durch Alanin ersetzt werden. Diese Methode wurde bereits früher verwendet, um die Peaks aus Driftzeitverteilungen bestimmten Konformationen der Prolyl-Peptidbindungen zuzuordnen. Wir zeigen hier, dass die Mutanten teilweise die Konformere des ursprünglichen Peptids wiedergeben, aber auch zusätzliche Strukturen bilden. Dies ist der höheren Flexibilität des Alaninstrangs im Vergleich mit Prolin geschuldet. Wir schliessen daraus, dass die Vertauschung von Prolin mit nicht-Prolin hilfreich bei der Zuordnung von Strukturen ist, jedoch in Verbindung mit spektroskopischen Techniken verwendet werden muss, da diese einen detaillierten Vergleich der Strukturen von mutiertem und nativem Peptid ermöglichen.

Diese Arbeit ist einer der ersten Schritte hin zum Aufbau einer Datenbank von Peptidstrukturen in der Gasphase mit atomarer Auflösung, sowie in Richtung eines verbesserten Verständnisses der Vorzüge und Grenzen von Spektroskopie und Ionenmobilität in der Gasphase, angewandt auf intrinsisch ungeordnete Peptide.

Schlagworte: Kryogenische Ionenspektroskopie, Feld-asymmetrische Ionenmobilitätsspektrometrie, Intrinsisch ungeordnetes Peptid, Cis/Trans Isomerisierung, Prolin, Isotopenmarkierung, Konformerselektive Spektroskopie, Kinetisch gefangen, Elektrospray, Bradykinin.

List of Abbreviations

BK   Bradykinin

BK[1-5]   Bradykinin 1-5 fragment

BNC   Berkeley Nucleonics Corporation

CC   Cis-cis

CCS   Collisional cross-section

CEM   Chain ejection model

CID   Collision-induced dissociation

CIS   Cold ion spectroscopy

CRM   Charged residue model

CT   Cis-trans

CV   Compensation voltage

DC   Direct current

DFT   Density functional theory

DT   Drift tube

DV   Dispersion voltage

EHSS   Exact hard sphere scattering

ESI   Electrospray ionization

FAIMS   Field asymmetric ion mobility spectrometry

FF   Force field

IEM   Ion evaporation model

IMS   Ion mobility spectrometry

IR   Infrared

IRLAPS   Infrared laser assisted photoelectron spectroscopy

IRMPD   Infrared multiphoton dissociation

ISC   Inter-system crossing

MBD   Many-body dispersion

MS	Mass spectrometry
NMR	Nuclear magnetic resonance
PA	Projection approximation
PCB	Printed circuit board
PDB	Protein Data Bank
RF	Radio frequency
SLIM	Structures for lossless ion manipulations
SRS	Stanford Research Systems
TC	Trans-cis
TM	Trajectory method
TOFMS	Time-of-flight mass spectrometry
TT	Trans-trans
UV	Ultraviolet

Table of content

Chapter 1.
Introduction

1.1 Peptides and proteins

Proteins make up around 15% of the mass of a typical eukaryotic cell and represent by far the largest class of biological molecules.[1] They take part in virtually all biological processes, serving as enzymes, signal molecules, receptors, ligand transporters and building blocks.[1] In recent years, significant progress has been made in understanding the mechanisms that allow proteins to perform their functions.

A poly-amino acid chain longer than 50-100 monomers is typically referred to as a protein, while shorter sequences are called peptides. The latter mainly act as signaling agents such as neurotransmitters, hormones, antibiotics and cell-penetrating agents.[2-10] They can be intentionally synthetized in the ribosomes as a part of a longer sequence and then tailored by special enzymes, or be produced as a byproduct of degradation of larger proteins. There is usually more than one receptor for a peptide family, and at the same time several peptides can bind to each receptor with different affinities, which ensures the flexibility of the regulatory system.[11] Given their length, peptides cannot form tertiary and quaternary structure, but they do adopt certain secondary structure, which is crucial for their interactions with larger proteins, DNA, and lipids.[9,10,12-15] The selectivity of their binding to other molecules is achieved *via* static or dynamic structural features.

Understanding the mechanisms of interaction between peptides and other biological molecules is of fundamental interest and has many practical implications, such as treatment of neurological diseases[16-19] and diabetes,[20] new antibiotics,[7,8,21,22] targeted drug delivery[12,23-25] and self-assembling materials.[26] The knowledge about peptides forms the basis for understanding the nature of larger proteins and their complexes on the molecular level.[27]

1.1.1 Traditional techniques for peptide structure determination

Structures of proteins and peptides are mainly determined by X-ray crystallography, NMR spectroscopy and, more recently, electron microscopy and mass-spectrometry.[28] Dominating secondary structure motifs can be also identified by circular dichroism. Here we briefly discuss advantages and limitations of the traditional solid- and liquid-phase techniques for peptide structure determination.

X-ray crystallography is the most common tool used to determine structures of proteins presented in the Protein Data Base.[29,30] It provides atomic resolution, and the measured parameters can be directly converted to structural information, avoiding molecular modeling.[31] The main disadvantage of this method is a need to crystallize the sample. Not all proteins form crystals, and the conditions required for crystallization are not predictable; thus up to 1000 sets of conditions might be tested for each system without guaranteed success.[31,32]

Even though X-ray chromatography is performed in the solid state, due to the presence of solvent molecules in the crystal, the degree of macromolecular crowding is comparable to that in the cell.[33] On the other hand, interaction between the proteins in the sample might alter their native conformations.[34] The latter is particularly true for peptides, which are small enough to be heavily influenced by the solvent and additives used for crystallization.[35] Cyclic mimics and conformationally restricted amino acids are often used instead of proteogenic sequences, which limits the applicability of the method.[36]

Nuclear magnetic resonance (NMR) spectroscopy can be performed directly in solution, and it is well suited for studying structure and, importantly, dynamics of small biological molecules.[37-39] NMR still lags behind X-ray crystallography in the number of structures submitted to the PDB due to the high costs and absence of an automated approach to structure determination[29] in spite of a large variety of 2D and 3D methods, which allow unambiguous assignment of the peaks and analysis of the interactions between different parts of the molecule.[40-42] One of the limitations of biological NMR is that measured parameters cannot be interpreted directly and structure determination requires molecular modeling.[43,44] In this work we use NMR to obtain additional information about the peptide in question in solution.

Circular dichroism (CD) is based on the principle that left-handed and right-handed circularly polarized light is absorbed differently by solutions of chiral molecules. CD allows fast and efficient secondary structure determination in terms of relative proportions of

alpha-helixes, beta-sheets, and random coils for a large range of protein sizes.[45] One of the main advantages of this method is that the sample remains in solution and can be studied under physiological conditions. Unfortunately, CD provides neither atomic resolution nor residue-specific information but rather an average signature of structural features.[46]

Cryogenic electron microscopy of biological samples has been rapidly developing in the last years.[47] Although resolution high enough to distinguish side chains and small molecules' complexes is achievable, this method is applicable for proteins and their complexes larger than 90 kDa, and thus structures of peptides are far beyond its scope.[48,49]

The techniques described above usually successfully determine the structure of biomolecules when they have well-defined secondary structure.[28] However, in the case of intrinsically disordered proteins and peptides (IDPs), NMR spectra often represent an average over an ensemble of different conformations.[50] Moreover, crystallization might not be possible due to the highly dynamic character of the disordered states.[39,51,52] That is why new conformer-selective approaches are needed to characterize structure, dynamics and function of IDPs.

1.2 Intrinsically disordered proteins and peptides

1.2.1 Natural abundance and biological functions of IDPs

The most widely used methods in structural biology have an intrinsic limitation: they overlook regions in the protein sequences that do not adopt well-defined secondary structure. Disorder predictors suggest that more than 50% of human proteins have long regions of disorder and 33% of eukaryotic proteins are mostly intrinsically disordered.[53,54] Such a state provides much greater variety of binding modes and kinetic regimes than a rigid protein structure.[54,55] IDPs play a key role in cell signaling, regulation and molecular recognition, in particular, in the nuclei.[54,56-62] Many questions remain regarding their nature: Does their flexibility come from solution conditions? What are the main factors contributing to the rapid structural changes? How many well-defined conformers are converting among one another? How does recognition by receptors or membrane insertion happen from a structural perspective? Is intrinsic disorder possible inside a membrane?[63]

1.2.2 Why IDPs in the gas phase?

Significant progress in the field of IDPs up to now has been achieved using NMR.[64,65] The main difficulty in the application of this method lies in the nature of the IDPs itself: in

solution they are found as an ensemble of quickly interconverting conformations, and any technique that seeks to provide insight into their structure has to separate signals coming from different conformers. This task can be dramatically simplified by the use of gas-phase conformer-selective techniques such as ion mobility and cold-ion spectroscopy coupled with mass-spectrometry.[66-69] Studying IDPs in the gas phase has a number of other advantages:

1. Profound understanding of their conformational space, including the height of the energy barriers, can be achieved, for example, *via* collision-induced isomerization or population transfer experiments.[70-76]

2. Some IDPs act in the environment of a cell membrane.[63,77] This medium can be reproduced in the gas phase using micelles.[78] Moreover, detergent release protects the structure of a membrane protein *via* competing with structural collapse.[79] Even if micelles are not employed, transition between water solution and the gas phase can, up to a certain extent, mimic the transition from the interior of a cell into its membrane.

3. The influence of solvent on the structure can be separated from the effect of primary structure. Solvent molecules can be added one by one and the occurring conformational changes can be observed.[80-86] Recently this method has been extended to lipids surrounding membrane proteins.[87,88]

4. Absence of solvent molecules simplifies or makes possible high-level quantum chemical calculations, which can be verified by direct comparison with gas-phase experiments.[89-91] This symbiosis of theory and experiment pushes forward the field of molecular modeling.[92]

5. Mass-spectrometry based methods have been quickly developing in recent years and now provide a wide range of experimental tools suited for structure determination.[93,94] Most studies of IDPs in the gas phase up to now were performed using IMS,[58,67,69,95,96] charge state distributions,[97] H/D exchange[98,99] or various MS/MS methods.[68,100] In this work we introduce the use of cold ion spectroscopy for characterization of IDPs, as discussed in section 1.6.

A major issue that prevents wider use of mass-spectrometry based techniques in structural biology is the relationship between the solution-phase and gas-phase structures or, in other words, the relevance of gas-phase structure for any biological process. This

question is especially important in case of IDPs since they do not have a well-defined native state.[101] Kinetic trapping seems to be a key to this issue.

1.3 Kinetic trapping

While Jarrold suggested in the 90's that mass-spectrometry "may provide a quick and sensitive tool for probing the solution phase conformations of biological molecules",[102] the range of validity of this statement is still under discussion. What is the relationship between the solution phase structure of a peptide and its gas-phase conformation? This question unavoidably arises whenever the results of mass-spectrometry based experiments are discussed.

The most widely used approach to transfer biological molecules from solution to the gas phase is electrospray ionization (ESI)[103-105] and by its nature it has the potential to preserve structural features of biomolecules. Despite years of research, the details of the electrospray process itself remain somewhat vague. Below we provide a brief overview of existing models and observations.

1.3.1 ESI models

The beginning of the ESI process has been well characterized.[106] When voltage is applied to the electrospray needle, the surface of the liquid is enriched with positive charges. It leads to formation of a Taylor cone, which results from an interplay between the surface tension of the liquid and the columbic repulsion of the ions on its surface.[107] If the applied electric field is high enough, the cone becomes unstable and a jet of small droplets emerges from the cone tip. There are several forms of the jet that were characterized using fast time-lapsed imaging and monitoring of the ES current; the most stable and widely used one is the cone-jet mode.[108,109] The charged droplets evaporate and decrease in size until the moment when the columbic repulsion between the charges on the surface of the droplets overcomes the cohesive force of the surface tension, which is referred to as the Rayleigh limit.[110] The droplets remove the excessive charge much in the same way as the Taylor cone does, that is by forming a jet of small progeny droplets.[111-113] The loss of mass in such events is typically less than 1%, while the loss of charge is much larger, ~ 15-25%, which means that the mass change of the droplets is determined primarily by the evaporation process. The evaporation/fission events iterate until nanoscale droplets are formed. In the nanospray setup the nanodroplet regime is reached faster, which decreases

the effect of background electrolytes.[114] The following models describe the last stage of the ESI process, which is similar for the nanoelectrospay and the conventional electrospray:[106,115]

- According to the *ion evaporation model (IEM)*, the biological ions of small to medium size can be ejected from the surface of a nanodroplet.[116,117] This is a spontaneous process with an activation barrier, which arises from an interplay between two opposing electrostatic forces: the repulsion of the ion by the charges on the droplet surface and the polarizability of the solvent.[118] In other words, in bulk polar liquid the ion prefers to be solvated, while as the charge density on the droplet increases, it becomes energetically favorable for an ion to evaporate.

- The *charged residue model (CRM)* assumes a large analyte ion, which stays inside the droplet until it fully dries out.[119] The charge of the final ion is in this case determined exclusively by its size and corresponds to the maximum charge that it can carry according to the Rayleigh limit.[120-122]

- The third mechanism, called *chain ejection model (CEM)*, is particularly important for disordered peptides and proteins: the ion extrudes from the droplet surface.[115,122,123] In general, this overlaps with the IEM, but since the biological ion is large, the transition state is more complex than originally considered in the IEM, and the evaporation process happens through a series of intermediate states. This model can explain why disordered proteins usually form higher charge states than their folded analogs.

Naturally, the ESI process plays a key role in shaping the structure of an ion in the gas phase. Usually, preservation of the solution-phase structure is associated with the CRM, which certainly prevails for large complexes in native-like conditions.[106,115,124,125] In the intermediate size regime, on which we focus in this work, a combination of all three mechanisms can be involved. We show that even for a small disordered peptide certain characteristics of its structure can be preserved, but additional studies are required to attribute these findings to a particular ESI mechanism.

1.3.2 Evidences of kinetic trapping

There are several types of evidence that some features of solution-phase conformation of a peptide or protein can be preserved when the molecules are transferred to the gas phase *via* electrospray:

- Collisional cross-sections (CCS) of the ions sprayed from native conditions are often close to those calculated for the crystal structure.[126-128]

- Infrared spectra of electrosprayed proteins in low charge states contain amide I and amide II bands typical for condensed-phase species.[129,130]

- Ions produced from denaturing and native conditions exhibit different properties, such as their charge state distribution, CCS, degree of gas-phase H/D exchange, or fragmentation patterns.[131-135]

- Proteins that are disordered in solution are observed in a wide range of conformations directly after the ESI process, while folded proteins that are electrosprayed under the same conditions are compact and exhibit a small range of charge states.[68,136]

- Gas-phase biomolecular ions have been shown to undergo structural changes (i.e., unfolding and re-folding) while trapped in a mass-spectrometer[137,138] or upon collisional activation.[139-141]

These experimental findings suggest that one can produce and preserve metastable species that are significantly different from the lowest energy gas-phase structures and are kinetically trapped, with barriers on the potential energy surface that inhibit them from isomerizing. Using cryogenic ion-mobility, Silveira and co-workers have recently provided evidence of kinetic trapping of triply protonated substance P in the gas-phase on the time scale of milliseconds after de-hydration.[82,142,143]

Determining whether solution-phase structures can be trapped and studied in the gas phase is important for evaluating the relevance of powerful, conformer-specific gas-phase techniques for biological applications. If significant structural elements can be preserved in the gas phase, one could selectively investigate the structure and dynamics of pre-selected, biologically active conformations, which is extremely difficult to do in solution. Understanding when electrosprayed ions are kinetically trapped is also important for the use of gas-phase experiments as benchmarks for theoretical predictions of structure. Since

theoretical methods rely on a global-minimum energy search for finding the most stable conformations, it may be difficult to find the structures that are produced in the gas phase if they are trapped in conformations that reflect solution conditions.

The idea of kinetic trapping forms the basis for the actively developing field of native mass-spectrometry, which seeks to utilize the power of gas-phase methods for structural biology.[94,106-109] This field is being developed mainly for large biological molecules and clusters.[125,144-147] Here we approach the issue of kinetic trapping from the other extreme and attempt to identify the smallest conformational features that can be preserved while transferring the ions from solution to the gas phase.

1.4 Proline *cis-trans* isomerization

The condition for kinetic trapping to occur is the existence of barriers on the free energy surface that are higher than the energy that the ions acquire while being transferred to the high-vacuum part of the instrument. Such barriers can occur due to *cis-trans* isomerization (Figure 1-1) as well as from a network of hydrogen bonds. For most natural amino acids the *trans* state of a peptide bond lies significantly lower in energy than the *cis* state and, accordingly, only 0.3% of all peptide bonds in the Protein Data Base are *cis*.[148] For proline residues, however, this number is estimated to be more than 5%.[148] Due to the cyclic nature of the proline side chain, the energy difference between *trans* and *cis* states is less than 2 kJ/mol in solution, but the barrier between them is high ($\sim$ 80 kJ/mol).[149]
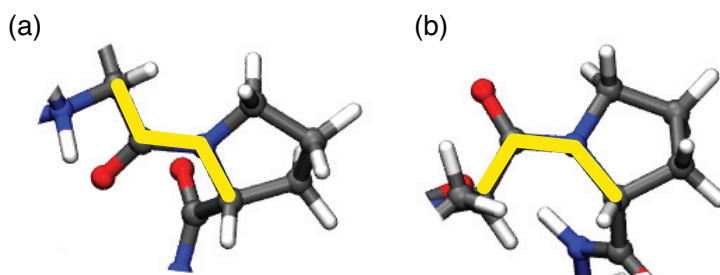

Figure 1-1. (a) *Trans* and (b) *cis* peptide bonds preceding a proline residue.

1.4.1 Proline in solution

Proline acts as a molecular switch,[150-154] and its isomerization is often a rate-limiting step in protein folding.[151,155-157] The structure of proteins is influenced by the presence of proline in the sequence in two somewhat orthogonal ways. On one hand, it disrupts helix

and beta-sheet formation and increases the level of backbone disorder.[158] On the other hand, proline is conformationally restricted and makes the protein backbone more rigid;[159,160] it also facilitates formation of certain types of β-turns.[161] We discuss the interplay between these two factors in detail in Chapter 5.

Intrinsically disordered parts of proteins and peptides are often proline-rich,[162] thus the question of proline conformation is particularly relevant to the structure and function of IDPs. Proline-rich motifs form so-called polyproline helixes, which serve as a unique binding interface or separate functionally different regions of IDP sequence. Single proline residues prevent formation of alpha-helixes or beta-sheets ("breaker" effect), which, among other functions, reduces the likelihood of spontaneous aggregation of IDPs.[163,164] In neuropeptides, proline is often found in the penultimate position, which facilitates formation of the *cis* isomer.[165]

1.4.2 *Proline in the gas phase*

Proline-containing peptides exhibit special properties that can be revealed by gas-phase techniques. Fragment ions produced N-terminal to proline are largely over-represented in collision-induced dissociation spectra.[166,167] Unusual fragmentation patterns are formed by proline-containing precursors upon photofragmentation using vacuum ultraviolet[168,169] and infrared multiphoton dissociation.[170-172] Furthermore, proline is one of the reasons for multiple resolved peaks in the ion mobility distributions of short peptides.[165,173-175]

In the gas phase the barrier for *cis-trans* isomerization is suggested to be significantly lower than in solution,[74] but it remains among the highest on the free energy surface of a peptide. This suggests that certain proline conformations can be kinetically trapped in the gas phase.[176]

1.5 Bradykinin as a model peptide

Getting a full dynamic picture of peptide conformations and their interaction with the environment is a complex problem. Generally, two approaches should to be applied in parallel: top-down, which starts from the average information about large parts of a biological system, and bottom-up, which studies in detail individual building blocks of increasing complexity. In our laboratory we follow the latter approach: we investigate

relatively short peptides in order to precisely determine their structure and to describe the factors that influence its formation.

For this study we chose a peptide that has been characterized previously using ion mobility spectrometry, but its structure in the gas-phase has not yet been determined. Bradykinin (BK, Figure 1-2) consists of nine amino acids of which three are prolines. It also has two phenylalanine residues, which makes it a convenient target for UV spectroscopy. The size of this peptide is at the limit of modern capabilities of high-level first principles structure calculations.
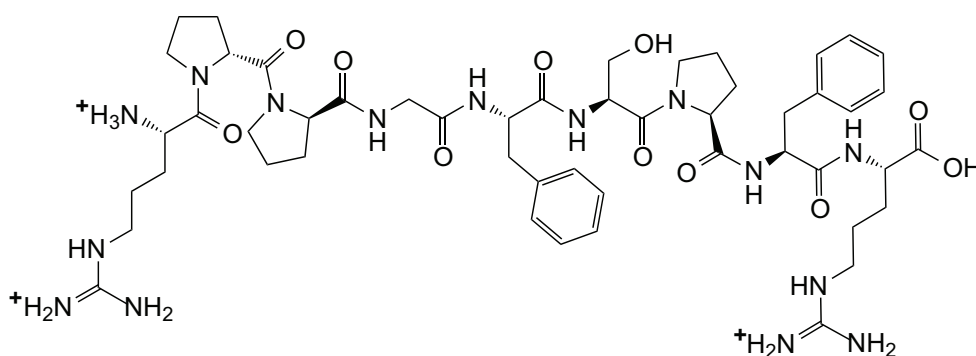
Figure 1-2. Chemical structure of triply protonated bradykinin.

Bradykinin plays a regulatory role in the cardiovascular and nervous systems and is a key reporter molecule in inflammation and pain.[177] It mainly acts through two receptors, namely, bradykinin B1 and bradykinin B2, which interact with the whole kinin peptide family.[11,178]

1.5.1   *Bradykinin structure in solution*

Already in the 70s the first NMR spectra of BK in aqueous solution were recorded.[179] They showed that in more than 90 % of the molecules all proline residues are in the *trans* conformation, which has been confirmed by later studies in water and in all other tested media including micelles.[180-183] The same conclusion was made for the complex of BK with bradykinin B2 receptor.[184,185]

In contrast to proline conformation, other structural motifs in the molecule are influenced by the environment. For example, in aqueous solution, no persistent structural features were identified, since the peptide rapidly interconverts among many conformers.[179] When certain polyphenols are added to the solution, BK can adopt a large

flexible turn between residues 6 and 9.[182,186,187] It is consistent with the notion that entering the membrane phase is often coupled with folding. The C-terminus forms a β-turn in most non-aqueous solvents, which suggests that this structural feature is important for peptide function.[180,188] Finally, upon binding to a receptor the peptide adopts a well-defined S-shaped structure.[184,185]

The main driving force for peptide folding in different non-aqueous solvents is the loss of intermolecular hydrogen bonds, which is similar to the transition to the gas phase.

1.5.2   Bradykinin structure in the gas phase

Over the years BK has been studied extensively, and presently it is often used for calibration of the instruments.[189] However, the exact structure of this peptide has not yet been determined despite numerous experimental and theoretical studies, which we briefly summarize here.

At first, BK was cationized using matrix assisted laser induced ionization (MALDI), which produces mainly singly charged ions.[190] The peptide was found to adopt a globular conformation independent of the cation used or of the temperature. When ESI became available, field-asymmetric ion mobility spectrometry (FAIMS) was combined with H/D exchange to study the most abundant +2 charge state, and four conformers were identified.[191] A theoretical study by Michael Siu and co-workers[192] attempted to find the lowest-energy structures of all charge states of BK in the gas phase. They concluded that the +2 charge state has a globular shape and prefers a zwitterionic form, while the lowest-energy structures of the +3 charge state are elongated and non-zwitterionic. The collisional cross-sections calculated by Siu were in agreement with the ones measured previously.[193,194] However, later Clemmer and co-workers showed that the +3 charge state of BK is protonated at two arginine side chains and the N-terminus, which is different from the theoretical prediction.[173] This finding puts into question the structures proposed by Siu for the triply protonated BK.

Bradykinin ions produced in the electrospray process are kinetically trapped and convert to a distribution of lowest-energy structures, originally denoted as a quasi-equilibrium distribution, upon collisional activation.[139] Remarkably, this is true both for the most abundant +2 charge state[195] and for the less populated +3 charge state.[139] On the other hand, Russell and co-workers report that formation of +3 charge state requires high temperature of the inlet capillary and is related to the desolvation thermodynamics.[82] They

also observe that doubly protonated BK is formed at low capillary temperature and does not undergo significant structural rearrangements (within the resolution of the instrument, R~30) during its slow desolvation.

Clemmer and co-workers suggested that the energy barriers between different conformational families of triply protonated BK come from *cis-trans* isomerization of prolyl peptide bonds[173] and used a sophisticated calibration procedure to estimate the height of these barriers.[74] They observe up to 10 features in the IM distribution under different conditions,[196] out of which the three main ones were assigned to *cis-cis-cis*, *cis-trans-trans* and *trans-trans-cis* conformations of the backbone. A detailed study using low-energy CID and IRMPD reveals a minor structural family within the most abundant peak in the IMS distribution, which was tentatively assigned to *cis*-Pro^2 and *trans*-Pro^7, i.e. opposite from the major conformational family.[70]

To summarize, bradykinin is well-suited system for a detailed study of the mechanisms of kinetic trapping. It is intrinsically disordered in solution and exhibits a variety of conformations in the gas phase for both charge states. By choosing a molecule that has been measured extensively in solution and in the gas phase, we seek to build a bridge between this two phases.

1.5.3 Bradykinin 1-5 fragment as a first step

Bradykinin is a nonapeptide, which is in principle accessible for modern first principles simulations. However, as shown in sections 1.5.1 and 1.5.2, it adopts a wide range of conformations, some of which are kinetically trapped. Identifying such structures theoretically is not trivial, since most methods of conformational search are intended to converge to the lowest-energy geometry. In order to establish an appropriate workflow for the simulations we had to choose a smaller test system that would exhibit similar behavior to BK. There is also a lower limit for the size of the peptide; since a sufficient number of hydrogen bonds is needed to stabilize the structure, kinetic trapping for di- or tri- peptides is hardly possible.[197]

From a functional perspective, biologically BK is inherently connected to the range of its homologues.[177] In the human body it is rapidly degraded by a number of enzymes, and the reported half-life of BK is 17 s.[198] The kinases cleave BK at 5 different sites, and some of the resulting fragments do not degrade further. In the first part of our study we focus on

the 1-5 fragment of BK (BK[1-5], Figure 1-3). There are several reasons to choose this particular fragment:

- This molecule is metabolically stable[198] and has an important biological role: it was shown to inhibit the enzymatic activity of thrombin.[199-201]

- BK[1-5] is significantly smaller than BK, which allowed us to develop and adjust an appropriate computational scheme.

- Residues 1-5 of BK do not adopt any secondary structure in all solvents investigated by NMR and represent the intrinsically disordered part of the sequence.
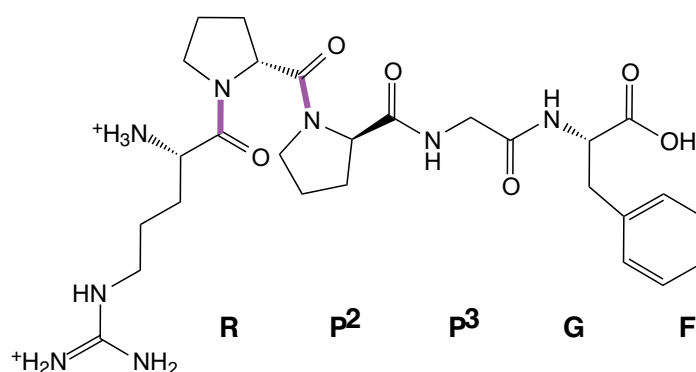
Figure 1-3. Chemical structure of doubly protonated bradykinin residues 1 to 5 (BK[1-5]$^{2+}$). The bonds around which the *cis-trans* isomerization can occur are highlighted.

Russell and co-workers[202] have investigated the +1 charge state of BK[1-5] by IMS and H/D exchange after MALDI. Three distinct structural forms were found, and their relative abundance depends on the composition of the solvent used to prepare MALDI samples. The extended form was observed when a high water concentration was used.

1.6 Conformer-selective gas-phase techniques

One of the main advantages of gas-phase techniques is the ability to separate ions based on a particular property, such as their collisional cross-section (CCS), mass-to charge ratio (m/z), absorption spectrum, the shape of charge-state distributions, fragmentation pattern, etc., and to draw conclusions about each sub-population of the ions separately. The techniques differ, however, by the level of detail they can provide. IMS, for example, allows one to measure only an average CCS of a conformational family,[203] while tandem MS

provides only indirect information about the structure. Infrared spectroscopy, on the other hand, gives access to the structure *via* comparison with computed vibrational spectra of a molecule.[204]

Charge-state-resolved infrared multiphoton dissociation spectroscopy (IRMPD)[205-208] and H_2-tagging experiments[209] have been used to determine the structure of many small molecules, clusters and complexes. However, when the systems under study get larger, their spectra become more congested because of overlap from different conformers.[210] To overcome this limitation we employ IR-UV double resonance spectroscopy, which allows one to obtain conformer-selective IR signatures.[211] This technique, in conjunction with first-principles simulations, has proven to be a powerful tool for atomic-resolution structure determination in the gas phase.[212-215]

Ultraviolet-infrared double-resonance spectroscopy requires a sharp electronic spectrum, which in turn requires cooling molecules to low temperatures.[216] The UV photofragmentation spectrum of a molecule is a sum of contributions from all conformers. Suppose that we increase the size of the system or study a structurally diverse molecule. The UV photofragmentation spectrum becomes congested, making the use of double-resonance technique challenging. This consideration led us to add another step of conformer separation prior to introducing the ions into the cold ion spectrometry instrument. We use drift-tube or field-asymmetric ion mobility spectrometry devices to physically separate the ions of significantly different conformations and apply either IR-UV double-resonance or H_2-tagging spectroscopy to each of the conformational families separately.

Combination of IMS with spectroscopy has already provided some important insights.[195,217-219] Here we apply it to model intrinsically disordered peptides in order to gain information about their structure on the way from solution to the gas phase.

1.7 Outline of the thesis

This thesis is organized in the following way.

Chapter 2 introduces the experimental and computational techniques that we employed. We describe two home-built instruments and the basic principles of the experimental methods, such as drift tube IMS, field asymmetric IMS, double-resonance cold ion spectroscopy and H_2-tagging spectroscopy. We also provide a brief explanation of the NMR measurements. From the first-principles simulations side, we describe all methods in

order of increasing complexity: force-field simulations and DFT calculations using in-house software FHI-aims.

Chapter 3 deals with the bradykinin 1-5 fragment. We first describe the experimental results obtained using a combination of drift tube IMS, cold ion spectroscopy and NMR, and then show how these results are used to guide and verify the theoretical conformational search. In this way we are able to identify not only the low-energy conformations of BK[1-5]$^{2+}$ in the gas phase, but also the kinetically trapped ones. We discuss the implications of these results for the connection between the gas and solution phase studies. We also draw a parallel between energy landscape of a molecule in the gas phase and that in the membrane or other non-polar media.

Chapter 4 shows the results of employing a similar set of experimental techniques to the full BK sequence. There is a vast amount of literature related to the NMR studies of BK, which we use for comparison, while in the gas phase apart from all methods listed in Chapter 3, we employ Field Asymmetric Ion Mobility Spectrometry (FAIMS) to introduce additional conformational selectivity. We show in detail how the combination of isotopic labeling of NH stretches, spectra of electronically excited state, and isotopic labeling of a chromophore allows us to assign the vibrational spectra of BK^{3+} and achieve a profound understanding of its conformational landscape in the gas phase.

In Chapter 5 we focus on the proline *cis-trans* isomerization in the BK sequence. We address the question whether a proline to alanine substitution can be used to identify conformational preferences of proline in the gas phase. To that end, we compare the conformational space of native BK with peptides where prolines in positions 2, 3 and 7 are substituted by alanines.

Chapter 6 summarizes the results and suggests directions for future research.

Chapter 2.
Experimental and computational approach

The main focus of this thesis is the gas-phase experimental study of peptide conformations. We use a multidimensional approach to identify unique structures of biomolecules and fully characterize them, that is, we separate the structural families using ion-mobility techniques and study them spectroscopically. From the spectroscopic standpoint, we have adopted two methods: tagging/messenger and double-resonance IR-UV spectroscopy. The former allows us to record a vibrational signature of all structures present in the ion trap at once, while the benefit of double-resonance spectroscopy is the ability to record conformer-selective vibrational spectra. We take advantage of both techniques, which are implemented in two separate home-built spectrometers.

As stated above, we use IM techniques to separate the ions prior to their injection into the spectroscopic part of the instrument, based either on their collisional cross-section in case of DT-IMS or the difference in their mobilities in the low and high electric fields, in case of FAIMS. Each of these IM techniques is used in one of the home-built machines. On top of that, we use mass-spectrometry based methods when appropriate, e.g. to characterize the fragmentation pattern.

A combination of all listed measurements creates a multidimensional space in which every conformation is characterized by a set of properties, such as CCS, UV and IR absorption, fragmentation pattern, etc. We then compare these data with the results of quantum-chemical simulations; moreover, we integrate the experimental information into the calculations.

Having explored the conformational space of the molecule in the gas phase, we then compare it to its behavior in solution. If the data are not available in the literature, we perform the NMR measurements and analyze them.

In the following sections we describe all mentioned techniques in detail. The information about gas-phase experimental methods is organized according to the instrument in which they are implemented. The general principles of each method are also presented. Separate sections are dedicated to the first-principles simulations and NMR spectroscopy.

2.1 Two complementary home-built spectrometers

The experimental part of this work is performed using two home-built instruments, which are complementary in many ways. The first one is a cold-ion spectrometer (CIS) equipped with a cryogenic octopole ion trap and two quadrupole mass filters. We perform double-resonance IR-UV spectroscopy on this machine, although tagging spectroscopy is also possible. A field asymmetric ion mobility spectrometry device (FAIMS) can be installed directly after the ion source for additional conformational selectivity. The second instrument combines a linear drift tube with a cold ion trap and a time-of-flight (TOF) mass spectrometer for measuring collisional cross-section distributions and vibrational spectra of mobility-selected ions. We apply hydrogen tagging to acquire vibrational spectra in this instrument. Below we describe the salient features of both spectrometers and emphasize the ways to decrease the ion collisional activation in light of the idea to preserve kinetically trapped species and to perform spectroscopy of them.

2.1.1 Cold-ion spectroscopy instrument

Using the cold ion spectroscopy instrument[220] (referred to as the CIS instrument) we have the ability to separate conformers using IR-UV double-resonance[216] in a cryogenic octopole ion trap and obtain electronic and vibrational spectra for each conformation. A schematic of the main parts of the CIS instrument is shown of Figure 2-1. The ions are produced in a nanospray source, and after passing through a metal capillary, which is usually kept at small positive voltage, they are injected into an ion funnel.
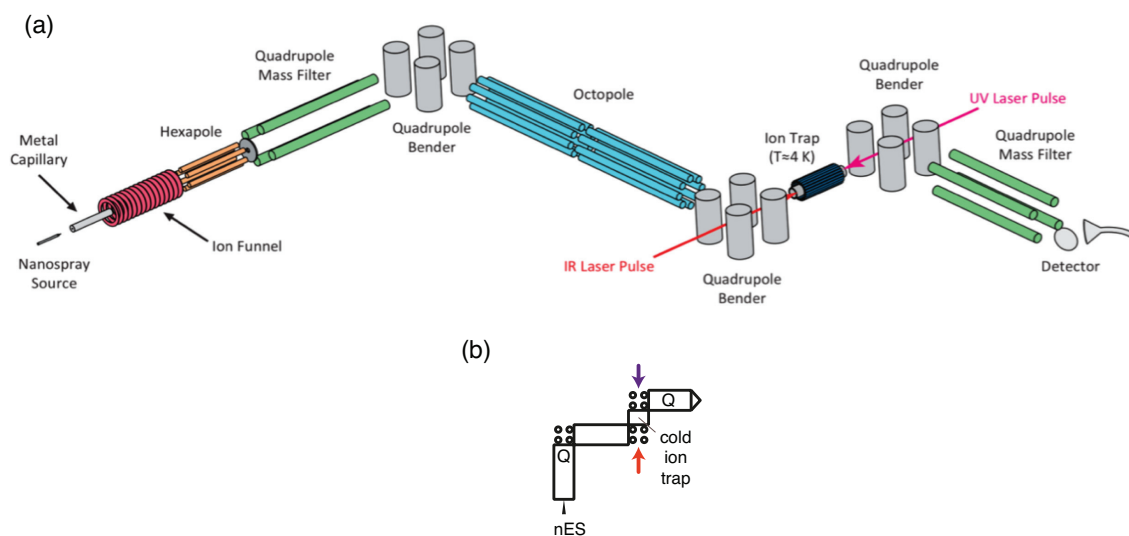
Figure 2-1. (a) Schematic of the cold ion spectrometer layout. (b) Block diagram used in the further chapters to represent this instrument.

The design of the ion funnel is based on the studies performed by Smith and co-workers.[221] It is an RF guide, which consists of 100 concentric plates connected through a resistor chain; the aperture of the last 42 electrodes gradually decreases in diameter. About 2 cm from the entrance a small conductive plate is suspended in the center of the funnel cross-section, which disrupts the directed jet of neutrals and is thus called a jet disrupter. An RF potential is superimposed on a DC gradient along the axis. The ion funnel improves signal stability and increases the number of ions by focusing the divergent ion beam coming out of the capillary. The downside of using the funnel is that some collisional activation of the ions invariably occurs at the end, where the aperture of the electrodes is the smallest.[222] When required, activation can be reduced by optimizing the DC voltage gradient over the length of the funnel, decreasing the RF amplitude, and decreasing the DC voltages applied to the last two electrodes. The optimal DC gradient is a compromise between two effects: if it is too high, the ions accelerate and experience energetic collisions, but if it is too low, the ions spend a long time thermally drifting through the ion funnel in the activating RF field. The conductance limiting electrode at the end of the ion funnel as well as the jet disruptor carry only DC potential.

The process of ion extraction from the funnel may collisionally activate the ions, since the background pressure decreases from 1.9 mbar to 0.04 mbar, which significantly increases the mean free path of the molecules:

$$\bar{l} = \frac{k_B T}{\sqrt{2} \cdot \sigma \cdot p},$$

(2.1)

where l is the mean free path, T – temperature, σ – cross-section of the molecules, p – gas pressure. Collisional energy in the center of mass reference frame is proportional to the electric field applied (E) and to the mean free path:[223]

$$\overline{K} \approx \frac{m_{gas}}{M} \left| \vec{E} \right| \cdot z \cdot \bar{l},$$

(2.2)

where m_{gas} is the molecular mass of the background gas, M – of the ion, z – charge of the ion.

To deal with this issue, the voltage drop between the end of the ion funnel and the pole bias of the hexapole is kept as low as possible. After being extracted from the funnel, the ions are pre-stored in a room-temperature hexapole ion trap. This allows thermal equilibration of the ions at 300 K and defines their reference potential, which corresponds to the pole bias of the hexapole. The ions are then ejected from the hexapole, mass-selected by a quadrupole mass-filter, and guided into a cold (4 K) octopole ion trap, where they are cooled in collisions with helium.[224] The kinetic energy of the ions entering the octopole trap is determined by the difference between the octopole pole bias and that of the room-temperature hexapole pre-trap. To avoid collisional activation this difference is kept small, typically less than 2 V. We do not observe any additional RF heating in the cold octopole,[225] and changing the RF amplitude does not seem to warm the ions.

As discussed in the following section, a set of laser pulses dissociates a fraction of the parent molecules. All the ions are then ejected from the trap, mass-selected by a second quadrupole, and detected by a channeltron.

2.1.2 *Principles of IR-UV double resonance spectroscopy assisted by IRMPE*

The typical spectroscopic scheme of the IR-UV double-resonance experiment is the following. The ions are cooled down to ~10 K in collisions with cold helium in the octopole ion trap.[225] After ~60 ms they are interrogated with a UV laser pulse, which promotes the ions to the electronically excited state causing a fraction of the parent molecules to fragment (Figure 2-2). Detecting the number of fragment ions as a function of the laser wavelength produces an electronic photofragmentation spectrum of the parent molecule. This spectrum is a sum of spectroscopic signatures of all conformers present in the trap. To obtain conformer-selective IR spectra of the cold ions, we tune the UV laser wavelength to a

specific transition in the electronic spectrum and introduce an IR laser pulse 200 ns before the UV pulse. This results in depletion of the photofragment signal, which is recorded as a function of IR wavelength.[216]
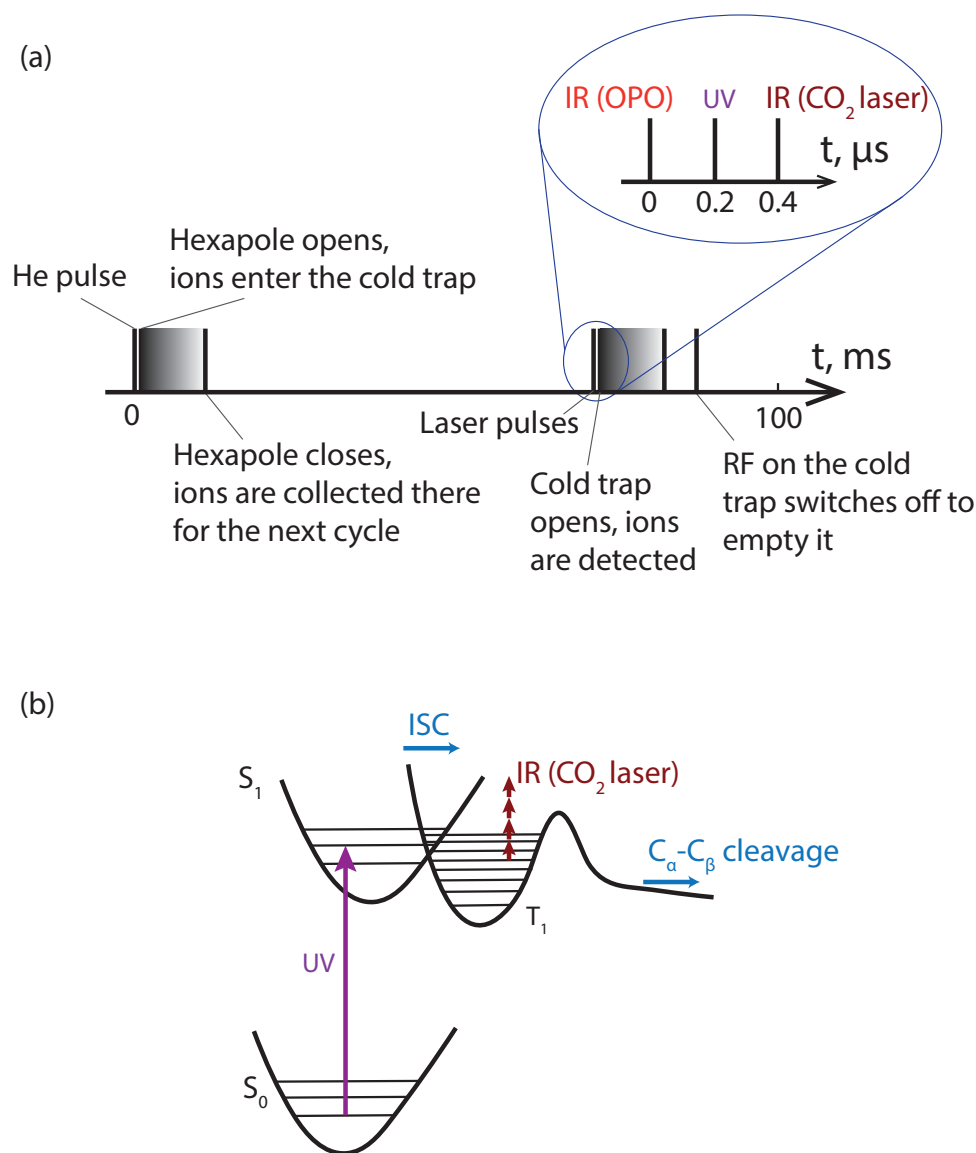
Figure 2-2. (a) Typical timing of the events within one cycle of the CIS instrument working at 10 Hz. (b) A model explaining infrared laser assisted photofragmentation of phenylalanine containing peptides.

Hole-burning double resonance spectroscopy allows us to record the vibrational spectra of the ions in the electronic ground state. In order to record the vibrational spectra of ions in the electronically excited state, we reverse the order of the laser pulses. First, a UV laser pulse promotes the ions to the S_1 excited state, and then an IR pulse excites a

vibration of the molecule within S_1, which increases the fragmentation yield. The delay between the pulses is kept as short as possible ($\sim$ 5 ns) to ensure that a large fraction of the ions remain in the S_1 state and do not have enough time to undergo neither ISC nor internal conversion. In this case the ions are vibrationally cold, and the IR laser excites them resonantly. Recording the gain in the photofragmentation as a function of the IR laser wavelength produces a conformer-specific vibrational spectrum of the electronically excited ions.[226,227]

The IR-UV double-resonance spectroscopic scheme described above produces relatively high photofragmentation signal for small phenylalanine- and tyrosine-containing peptides. However, in the case of larger phenylalanine-containing peptides, the fragmentation yield is often low and, and as a consequence, the signal-to-noise ratio is also low. In order to overcome this limitation we add a CO_2 laser pulse 200 ns after the UV laser to allow infrared multiphoton excitation (IRMPE) (Figure 2-2 (a)), which increases the fragmentation yield.[228] The photophysics of the UV excitation and the following processes has been the subject of a separate study,[226] which concluded that upon UV excitation a fraction of the parent ions undergo intersystem crossing to a long-lived triplet state, which is believed to have a barrier for dissociation of the $C_\alpha - C_\beta$ bond.[226] Adding a CO_2 laser pulse helps the ions to overcome this barrier, producing phenylalanine side chain loss fragments, which are not present in the statistical fragmentation pattern of vibrationally excited ground-state parent ions. This model is summarized in Figure 2-2 (b). In all experiments presented in this work we use side chain loss fragments to record the photofragmentation spectra.

The UV pulses are produced by doubling the output beam of a Lumonics tunable dye laser (HD-500) in a BBO crystal. An Inrad Autotracker-III maintains the optimal phase match angle as the laser is scanned. The dye laser, which is used with Coumarin 540A, is pumped by the third harmonic of the Nd:YAG fundamental (355 nm). The length of the UV pulse is $\sim$5 ns, while the line width of the dye laser is 0.07 cm^{-1}. This limits the resolution of the electronic spectra presented in this work. The IR laser pulses in the 3 μm region are produced by a tunable OPO (Laser Vision), which is pumped by the fundamental of a Nd:YAG laser. The length of the pulse is $\sim$10 ns and the line width 1.5 cm^{-1}. Finally, a Lumonics TEA-840 CO_2 laser is used to generate IR radiation in the 10 μm region. This pulse has an overall width of $\sim$5 μs consisting of an intense spike of 100-200 ns followed by

a longer, lower intensity tail. The power is kept in the range 300-600 mJ. This laser is line-tunable, and we use it on the brightest line at 10.6 µm.

2.1.3 Field asymmetric ion mobility spectrometry

The electronic spectrum measured by UV photofragmentation spectroscopy is a sum of individual photofragmentation spectra of each conformer present. For large molecules the spectrum becomes so congested that IR-UV double-resonance spectroscopy alone can no longer record a vibrational spectrum of an individual conformer. In order to push the size limit of the systems under study it is necessary to pre-filter the conformers before loading them into the cold ion trap. One of the ways to do this is to use field asymmetric ion mobility spectrometry (FAIMS). Below we briefly present the physical principles and implementation of FAIMS.

The mobility of an ion K in an electric field E is defined as a ratio between its drift velocity and the strength of the field:

$$K = \frac{v}{E}.$$

(2.3)

In the low-field regime the mobility is an intrinsic property of an ion and does not depend on the strength of the applied electric field, which gives rise to linear ion mobility methods, which we discuss in section 2.1.5. However, in high electric fields (more precisely, when the normalized field intensity E/N is high[229]) the mobility depends on the strength of the electric field, and different species can be separated using this dependence.

There are several effects, which give rise to such dependence.[229] *The standard high-field effect* is omnipresent and results from the fact that in the high electric field the velocity of the ions is determined only by the field and not by thermal motion. A simple calculation shows[229] that in this case the drift velocity of the ion is proportional to the $(E/N)^{1/2}$:

$$v = \xi^{1/2} \left(\frac{1}{\mu m}\right)^{1/4} \left(\frac{qE}{\Omega N}\right)^{1/2},$$

(2.4)

where ξ is a dimensionless parameter close to 1, m is the mass of the buffer gas, μ – reduced mass of the ion and the buffer gas molecules, q – the charge of the ion, Ω – collision integral (CCS), N – number density of gas molecules. It means that in very high electric fields the mobility of the ions K is:

$$K \sim E^{-1/2},$$

(2.5)

while in the low-field limit K is independent of E. FAIMS operates in an intermediate regime, where the dependence of the ion mobility on the electric field can have different shapes depending on the interacting potential between the buffer gas and the ion.[230]

Another effect that contributes to the difference in the mobility of the ions in high and low electric fields is a *clustering-declustering mechanism*, which is especially pronounced for polar buffer gases.[231-233] At low electric field, clustering of the ions with the buffer gas will increase the collisional cross-section. Under the influence of high electric field the temperature of the ion-buffer cluster increases, which causes its dissociation and decreases the effective cross-section of the ion.[229]

At low electric field the observed collisional cross section of the ions is averaged over all possible orientations. However, a high electric field can align the ions, producing a change in the observed mobility *via* two mechanisms: *collisional alignment and dipole alignment*.[234] In the case of high electric field the ions undergo more collisions in the direction of the drift than in perpendicular directions. This causes the ion to orientate itself minimizing the cross-section in the direction of the motion. On the other hand, if the ion has a high dipole moment, it orientates itself along the electric field lines in the high field, which causes a change in the observed mobility of the ion.[235] However, for room-temperature ions this effect is important only for dipole moments higher than 400 D, which is possible for the proteins heavier than 30 kDa.[234]

Practically, FAIMS is implemented in the following way. The ions produced in the electrospray source are driven by the carrier gas into a gap between two electrodes. An asymmetric waveform is applied to the electrodes such that the time-averaged voltage is zero. If the mobility of the ions were the same in high and low electric fields, all species would pass through. Since this is not generally the case, a small compensation voltage (CV) has to be superimposed on the waveform. Scanning the CV produces a distribution of ions as a function of the difference in their mobilities at high and low electric fields, which we denote a CV distribution.

Several different shapes of electrodes are used in FAIMS: flat,[236] cylindrical,[237] "dome"[238,239] or "side-to-side" geometries.[240] We have used the latter, shown in Figure 2-3. The temperature of the electrodes can be controlled independently using a flow of compressed air.[241] There are two reasons to employ active temperature control. First, it allows one to avoid slow drift of the CV because of FAIMS device heating caused by other actively heated elements of the instrument, such as a heated capillary.[242] Second, a temperature gradient

across the analytical gap has been shown to improve the resolution of FAIMS by compensating for the radial inhomogeneity of the field.[241] However, for the molecules in this study no significant improvement of the resolution was achieved by changing the temperature gradient. We do not use any heated elements in the ion source, thus we kept all electrodes at room temperature during all experiments.
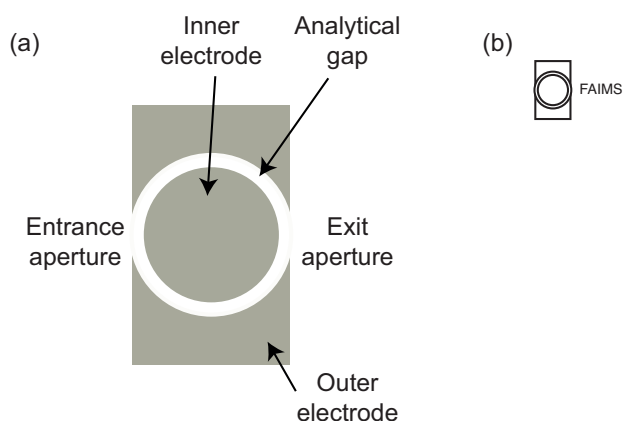
Figure 2-3. (a) Schematic of the FAIMS "side-to-side" geometry used in this work. (b) Drawing used in the further chapters to schematically represent FAIMS.

The optimal shape of the asymmetric waveform applied to the FAIMS electrodes depends on the exact shape of the *K(E)* dependence for an analyte.[243] In our study we use the most common bisinusoidal waveform,[244-246] shown in Figure 2-4.
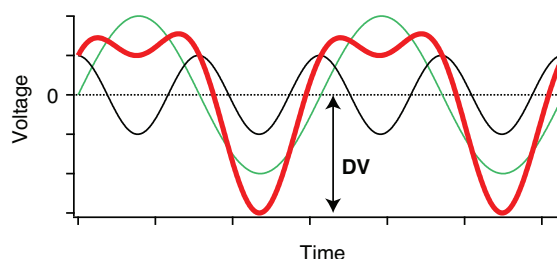
Figure 2-4. Typical asymmetric waveform applied to the electrodes of FAIMS (thick line) is a sum of two sinusoidal functions, shifted in phase by $\pi/2$ (thin lines).

The resolution of the FAIMS depends both on the amplitude of the asymmetric waveform (DV) and on its frequency. In this work we keep both constant: the DV is set to the highest possible value (5 kV) and the frequency of 750 kHz is determined by the construction of the waveform generator.[247] Other parameters that influence the resolution and the sensitivity of FAIMS include the carrier gas flow rate[248-250] and composition.[239,251-]

254 We find the optimum value of the carrier gas flow rate to be 2 l/min using N_2 as a carrier gas.

2.1.4   Combining FAIMS with Cold Ion Spectroscopy

Thermo Scientific has implemented the design described in paragraph 2.1.3 in a commercially available FAIMS interface, which we used without major modifications. It is compatible with the Thermo ion source, which makes combining FAIMS with the home-built CIS spectrometer, presented in paragraph 2.1.1, straightforward.

The tandem FAIMS-CIS instrument is used in three regimes:

- First, the CIS part of the instrument is used simply to record the CV distribution of the parent ions. In this mode the quadrupoles in the CIS machine are set to the mass-to-charge ratio of the parent ions, and all parameters are optimized to maximize their transmission.

- Second, FAIMS is used as a conformational filter for spectroscopy. We set the CV on a value corresponding to a given conformational family and use the cold ion trap to further investigate the ion structure *via* UV spectroscopy or IR-UV double resonance.

- In a third regime, cryogenic ion spectroscopy is used as a high-resolution, selective detector for FAIMS. We park the UV laser on a peak in the UV photofragmentation spectrum, which corresponds to a given conformer. The ions are exposed to a laser pulse every second trapping cycle. The first quadrupole is set to the mass of the parent ion, while the second quadrupole is switching between the mass of the parent and the mass of a photofragment. The CV is scanned and two CV distributions are recorded simultaneously: when the laser is off, a parent ion CV distribution is obtained, while when the laser is on, a CV distribution of only the chosen (by the UV laser) conformer is recorded. The measurement is then repeated for all major conformers and the CV distribution of the parent ion is decomposed into a sum of individual contributions of each conformer effectively increasing the resolution of FAIMS.

Examples of the data obtained in each of these modes are presented in subsequent chapters of this thesis.

2.1.5 Drift-tube ion mobility spectrometry

The main advantage of linear ion mobility compared to FAIMS is the possibility to extract structural information from the measured arrival-time distributions. This is achieved by operating in the low-field regime, so that the thermal energy is higher than the energy supplied by the electric field.

In the simplest configuration of drift-tube ion mobility, ions migrate the length of a tube filled with gas under the influence of a low electric field in the axial direction. The arrival time t_{arr} is measured, and since the distance, l, that the ions travel, is predetermined, the average drift velocity of the ions can be easily calculated:

$$v = {l}/{t_{arr}}. \tag{2.6}$$

In the low-field regime, the drift velocity is proportional to the applied electric field:

$$v = K \times E. \tag{2.7}$$

The proportionality constant K is related to the collisional cross section (CCS) of the ion:[255]

$$K = \left(\frac{3z}{16N}\right)\left(\frac{2\pi}{kT}\right)^{1/2}\left(\frac{m+M}{mM}\right)^{1/2}\left(\frac{1}{\Omega}\right), \tag{2.8}$$

where N is the number density of the buffer gas, z is the charge of the ion, k is the Boltzmann's constant, T is the absolute temperature, m is the mass of the buffer gas, M is the mass of the ion, and Ω is the spatially averaged collisional cross-section.
Together with

$$V = El, \tag{2.9}$$

$$p = NkT, \tag{2.10}$$

$$\mu = \frac{Mm}{M+m}, \tag{2.11}$$

where V is the voltage across the drift tube, p is the pressure of the buffer gas, μ is the reduced mass, this gives rise to

$$\Omega = \frac{t_{arr}V}{l^2}\left(\frac{3z}{16p}\right)\left(\frac{2\pi kT}{\mu}\right)^{1/2}. \tag{2.12}$$

Thus the CCS can be directly calculated from the arrival time of the ion using only experimentally measurable parameters.

The resolution of the DT-IMS is determined by the degree of broadening of the ion packet due to diffusion. It was shown[255] that

$$\frac{t_{arr}}{\Delta t_{arr}} = \left(\frac{Vz}{16kTln2}\right)^{1/2}.$$

(2.13)

It is obvious from this formula that the resolution of the linear DT-IMS instrument can be increased by increasing the length of the drift tube or decreasing its temperature.

Drift tube ion mobility is the first and the most common type of IMS.[256-259] A typical 1 m long drift tube reaches the resolution of ~70,[260] while the cyclotron-geometry drift tube developed in the group of Clemmer achieved R~300.[261] In recent years other ways of measuring CCS have been developed to improve the resolution and sensitivity.[260,262,263] While still based on the proportionality between the velocity of the ions and the applied electric field, they employ different strategies:

- *IM with radial collisional focusing* utilizes the middle quadrupole of a triple-Q system as a drift cell.[194] RF focusing ensures good ion transmission and increases the sensitivity. Even though the authors claim that the RF heating of the ions is negligible, this type of IMS has not been widely adopted.

- The design of the cell for *travelling wave IMS*[264-266] is similar to that of the traditional drift tube – a stack of ring electrodes – but the electric field is not homogeneous. The ion propagation is produced by superimposing a voltage pulse on a confining RF field and moving the pulse along the drift cell to provide a travelling wave, on which the ions can surf. This approach does not require high voltage across the drift tube and has been readily adopted in commercial instruments.[265] The drawback of such a scheme is that the CCS cannot be retrieved directly from the arrival times and thus calibration using known compounds has to be performed.[267-270] The group of R. Smith has recently developed structures for lossless ion manipulation (SLIM), which utilize the travelling waves, but the ion guide is produced using PCB technology.[271-273] Using multi-pass devices, they have achieved resolution of R~500.[274]

- In *Trapped IMS* the ions are held in place by a balance between the drag force coming from the flow of gas and an electric force.[275] The IMS instruments of this design are compact and achieve resolving power exceeding 300.[276]

- Within the *variable temperature DT-IMS*[277] the most prominent results were obtained using cryogenic IMS, which allows investigation of the sequential dehydration of the ions during the last step of the ESI process.[81,82,278]

- *Transversal Modulation IMS* produces a continuous flow of mobility-selected ions with a moderate resolving power ~55 using a geometry similar to FAIMS.[279,280]

- Finally, *tandem IMS*, first constructed by D. Clemmer,[281] combines two or more consecutive drift sections separated by selection gates, ion funnels and activation regions. A two-section version of a tandem IMS instrument is used for this study, and it is described in detail in the following paragraph. An ultimate version of tandem IMS is a circular drift tube, which is an elegant idea of extending the length of the drift region without increasing the physical size of the instrument.[261]

2.1.6   Tandem IMS-IMS instrument

The linear drift tube, which we use for our experiments, was originally constructed in the group of D. Clemmer[281] and then re-assembled in our laboratory to be used in conjunction with our CIS instrument, as described in section 2.1.10. Figure 2-5 gives an overview of the hybrid IMS-CIS instrument. In this section we focus on the design of the tandem IMS part.

The ions are produced in a home-made ESI ion source and driven through a metal capillary to the first ion funnel, which has an hourglass shape[282] and serves to accumulate the ions prior to their injection into the drift tube. The ion gate in the end of it is equipped with a mesh grid to prevent penetration of the RF field from the ion funnel into the drift region. The gate opens every 100 ms for 100 μs to initiate the process of ion separation. The initial width of the ion packet should be kept much shorter than the typical drift time, which reaches 5-10 ms for the size of the systems we are investigating in this work.

The RF amplitude applied to the first ion funnel significantly influences the level of ion activation and is kept as low as possible. Other parameters that influence the degree of activation include the voltage difference across the ion funnel, the voltage applied to the last electrode of the funnel relative to the trapping voltage, the trapping voltage V+ applied to the closed ion gate, and, most importantly, the voltage on the jet disrupter, which has to be kept as low as possible to avoid pushing the ions to the regions of high RF fields.
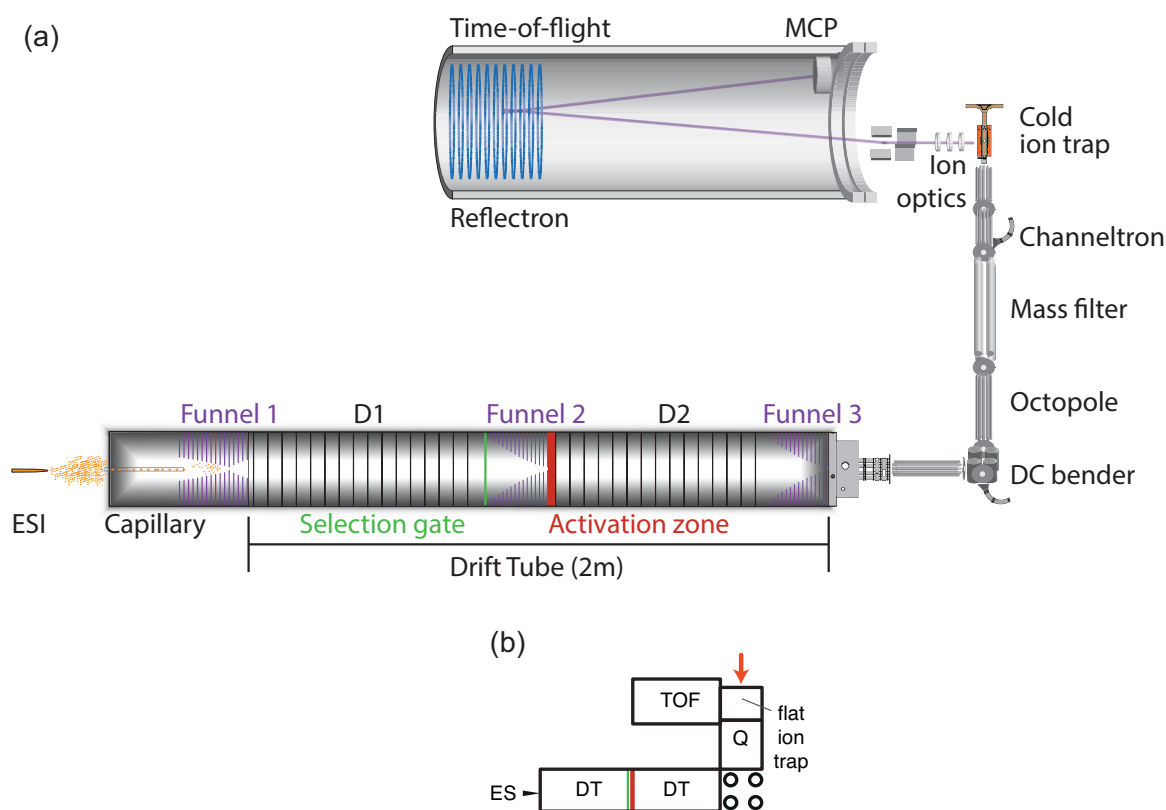
Figure 2-5. (a) Schematic of the hybrid IMS-CIS instrument. (b) Block diagram used in the further chapters to represent this instrument.

The ion packet initiated by the opening of the ion gate propagates through the drift tube filled with He at a pressure of ~3 mbar. A 1.6 m long tube consists of two separate drift regions of equal lengths. The electric field in the drift regions is kept constant at 11.3 V/cm, while in the second and third ion funnels it is elevated to 12.6 V/cm to insure good ion transmission and avoid trapping ions at the end of the middle ion funnel.[281]

If the instrument is operated as a single high-resolution drift tube, the ions are allowed to propagate without interruption until the last electrode of the third ion funnel, after which they enter a high-vacuum chamber. This transition point is a potentially activating region of the machine, and the voltages applied to the ion optics should be adjusted to insure low activation. Note that if the cold ion trap is not used for spectroscopy, such precaution is unnecessary.

A second ion gate is situated between the first drift region, D1, and the middle ion funnel and serves to "slice" out a portion of the ion mobility distribution produced in the first half of the drift tube. The timing of the gate is set using a Berkeley Nucleonics Corporation (BNC) pulse/delay generator (model 575), and the width of the selection pulse

is typically 100 µs. After passing through the middle ion funnel, the ions enter the activation region formed between the last electrode of the ion funnel and the first electrode of the second drift region, which are separated by 0.3 cm. The voltage across this region can be increased up to 140 V, and the drift time distribution of the activated ions is observed as a function of this voltage. Applying the maximum voltage of ~80-100 V usually produces fragmentation of the parent ions.

After being separated in the second drift region, D2, and passing through the third ion funnel, the ions enter a high-vacuum chamber and pass through two octopole ion guides and a quadrupole mass filter. The resulting arrival time distribution of mass-selected ions is recorded using a channeltron, which can be translated into the ion beam.

The DC voltages on the electrodes of the drift tube are produced using power supplies from Stanford Research Systems (SRS) and Agilent, connected in a chain so that every power supply is floated to the output voltage of the previous one. An example of such connection is shown in Figure 2-6 for the electrodes in the middle funnel region.
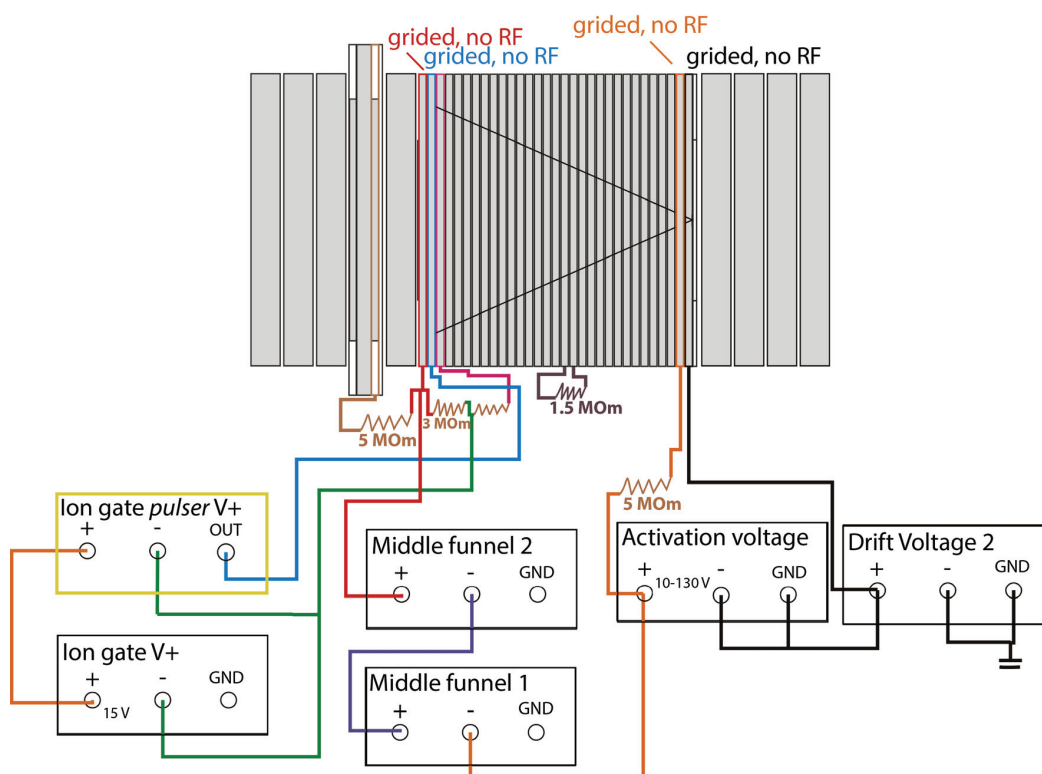

Figure 2-6. Schematic of the electrical connections between the first and the second drift regions. Note that the power supplies are connected in a chain, and each one is floated to the output voltage of the previous one.

2.1.7 Direct measurement of the CCS

The following measurement procedure is based on the protocols developed byK. Pagel and co-workers.[283]

There are two ways to convert the arrival time distribution measured in a linear drift tube to a CCS distribution. The first one is a direct method using formula 2.12. An obvious advantage of this method is that it does not require comparison with other measured CCSs and in this sense can be called *ab initio* CCS determination. It should be used when a compound of a new chemical class is investigated.

The main disadvantage of the direct measurement is that the arrival time t_{arr} corresponds to the time that the ion spends in the drift tube itself, while the measured time consists of the true arrival time t_{arr} plus any dead time that the ion spends in the high vacuum parts of the instrument:[269]

$$t_{measured} = t_{arr} + t_0$$

$$= \frac{\Omega l^2}{V}\left(\frac{3z}{16p}\right)^{-1}\left(\frac{2\pi kT}{\mu}\right)^{-1/2} + t_0. \qquad (2.14)$$

In order to find Ω we vary the voltage across the drift tube V, record the arrival times for ions of known mass and charge, and plot $t_{measured}$ as a function of *1/V*.

We performed a direct measurement of CCSs for the +2 and +3 charge states of bradykinin. The latter is separated into three well-resolved peaks, and the CCS of each of them was measured separately. The electrical field was kept uniform throughout the length of the drift tube, including the ion funnels, and its strength was varied from 6 to 11 V/cm. We measure the voltages across each of the drift regions separately, calculate the ratio *l²/V* for each segment, and then sum them up. Other parameters were kept constant: *p* = 3.65 mbar, T = 298.6 K. The peak positions in time were picked using a procedure built into Igor Pro 6.3. The resulting plots are shown in Figure 2-7, while the values, which were calculated using linear regression, are summarized in Table 2-1.

The directly measured CCSs differ from the ones measured previously by less than 2%,[139,193] which is in the range of random errors typically assumed for CCS measurements. The systematic error is typically smaller.
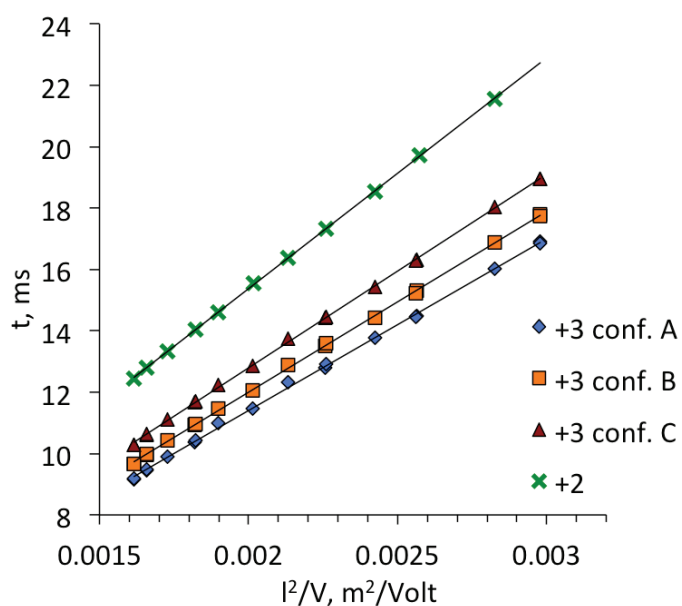
Figure 2-7. Dependence of the observed arrival time on l^2/V and linear regression for the three conformations of BK^{3+} (A, B and C) and BK^{2+}.

Table 2-1. Results of the direct measurement of the CCSs of doubly and triply protonated BK. t_0 is the time correction; "difference" corresponds to the difference between the CCSs measured here and the literature values.

species	M, a.u.	z	t_0, ms	CCS measured, Å^2	CCS literature, Å^2	difference, %
+3 conf. C	1063	3	0.177	307.6	304[139]	1.2
+3 conf. B	1063	3	0.217	287.3	285[139]	0.8
+3 conf. A	1063	3	0.209	272.9	269[139]	1.5
+2	1062	2	0.328	244.5	246[193]	0.6

The ratio between the "dead time" t_0 and the total drift time appears to be of the same order of magnitude as the statistical error, ~2%, which means that the "dead time" cannot be neglected. It depends on the mass and charge of the ion, but can be converted to a universal coefficient using a simple consideration. In the absence of the buffer gas the ions move with acceleration imposed by the electric field and their movement at each given moment can be described as

$$M\ddot{x} = Ez,$$

$$\frac{\ddot{x}\Delta t^2}{2} = \Delta x. \qquad (2.15)$$

Taking into account that the length of the path is approximately the same for any ion, we see that the time correction

$$t_0 \sim \sqrt{\frac{M}{z}}. \tag{2.16}$$

From this argument also follows that the "dead time" depends on the applied voltages and has to be determined for every tuning of the machine. We averaged the reduced time correction, independent of the ion mass and charge, for all rows in Table 2-1 and found that in our case it is

$$t' = t_0 \sqrt{\frac{z}{M}} = (1.2 \pm 0.3) \times 10^{-2}, \tag{2.17}$$

where M is expressed in a.u, z is dimensionless, and t_0 is in ms.

2.1.8   Calibration of the drift tube using BK

The second method of converting arrival time distributions to the CCSs is by calibration, using molecules of known average cross section, which in this work was bradykinin. First, the drift times of the calibrant and of the unknown ions are measured under the same conditions.

Note that the charge states and the masses of the set of calibrants and the unknown ions are likely to be different. The drift times of the calibrants are

$$t_{measured} = t_{arr} + t_0$$

$$= \frac{\Omega l^2}{V} \left(\frac{3z}{16p}\right)^{-1} \left(\frac{2\pi kT}{\mu}\right)^{-1/2} + t_0 \tag{2.18}$$

$$= C_1 \Omega \frac{\sqrt{\mu}}{z} + C_2 \sqrt{\frac{M}{z}},$$

where C_1 and C_2 depend only on the parameters of the instrument, but not on the masses and charge states of the ions. A linear regression is built based on the data acquired for the calibrants as:

$$t_{measured} \times \sqrt{\frac{z}{M}} = C_1 \Omega \sqrt{\frac{m}{(m+M) \times z}} + C_2, \tag{2.19}$$

where m is the molecular mass of the carrier gas. The linear fit produces the values of C_1 and C_2.

Computing the CCSs of the unknown system based on the arrival time distribution t is the reverse procedure:

$$\Omega_1 = \frac{1}{C_1} \sqrt{\frac{z_1(m + M_1)}{m}} \left(t \times \sqrt{\frac{z_1}{M_1}} - C_2 \right), \qquad (2.20)$$

where M_1 is the mass of the ion of interest and z_1 its charge.

2.1.9 Calibration of the drift tube for tandem IMS

The described calibration procedure was developed for the linear drift tube and TWIMS.[267] In the case of tandem IMS, when the ions are gated in the middle of the instrument, an extra calibration procedure is required. We tested which parameters change the observed arrival times of BK^{3+} and found that ion activation in the activation region does not (Figure 2-8). On the other hand, the selection event itself reduces the drift times by ~0.5%. This value should be compared to the difference in the CCSs between two conformers, which is 4% for the closest peaks. We see that the effect of selection cannot be ignored and re-calibration is needed when the ions are gated after the first drift region. We propose the following procedure to perform it.
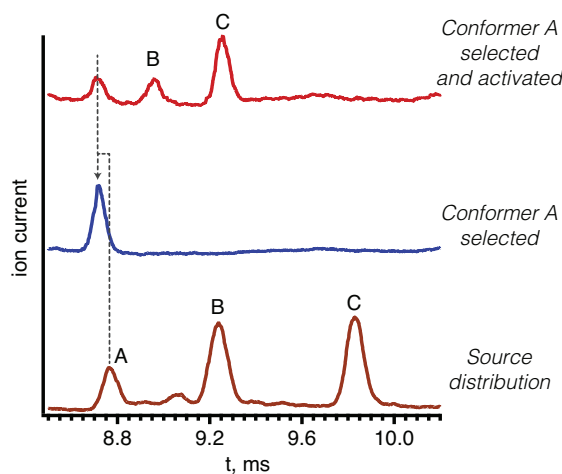
Figure 2-8. The measurements required for the calibration procedure for IMS-IMS. One of the conformers of BK^{3+} is selected and it results in a shift in the drift times towards shorter values. The activated distribution is obtained by applying 30 V in the activation region and the drift times for conformers B and C are retrieved.

From the drift time distribution of BK^{3+} produced in the first drift region we select peak A and activate it with ~30 V to partially convert it to conformational families B and C

(Figure 2-8). We assume that the CCSs of conformers A, B and C are already known. The fact that peak A partially converts to B after being selected and passing through the middle ion funnel can be described mathematically as:

$$t_{total}^{(A)} = t_{selection} + C_{if}\Omega'_A \frac{\sqrt{\mu}}{z} + C_2\Omega'_A \frac{\sqrt{\mu}}{z}.$$

$$t_{total}^{(B)} = t_{selection} + C_{if}\Omega'_A \frac{\sqrt{\mu}}{z} + C_2\Omega'_B \frac{\sqrt{\mu}}{z},$$

(2.21)

Here t_{total} is the corrected drift time for a given conformer with the selection switched on, $t_{selection}$ – the time when conformer A was gated at the end of the first drift region, Ω' – the reduced CCSs, and μ - the reduced mass of BK and He. The coefficients C_{if} and C_2 correspond to the middle ion funnel and to the second half of the drift tube and depend on the parameters of the instrument, but not on the nature of the ions. Subtracting one equation from another we get:

$$t_{total}^{(B)} - t_{total}^{(A)} = C_2(\Omega'_B - \Omega'_A)\frac{\sqrt{\mu}}{z}.$$

(2.22)

From this relation we find C_2, which allows us to convert the unknown activated drift time distribution to CCSs. Indeed, if we select a peak with a given CCS Ω'_1, which was found previously using standard IMS, these ions arrives with the corrected time t^1_{total}. Upon activation we get a corrected arrival time distribution t, which we convert to CCSs using the following expression:

$$t - t_{total}^1 = C_2(\Omega' - \Omega'_1)\frac{\sqrt{\mu_1}}{z_1}.$$

(2.23)

This calibration procedure can be extended to other calibrants if molecules of significantly different sizes are measured. For the systems discussed here, the potential uncertainty coming from using only three peaks of BK for the calibration appears to be lower than the random error, which we found to be on the order of 1% (see paragraph 2.1.7).

2.1.10 Hybrid IMS – CIS instrument

We combined tandem IMS with cryogenic spectroscopy in order to obtain vibrational signatures of mobility-selected ions. The schematic of the hybrid IMS-CIS instrument is shown in Figure 2-5. We have described the first drift-tube part on this machine in section

2.1.6; we now discuss the cold ion trap used for spectroscopy as well as the TOFMS used for detection.

The CIS instrument described in section 2.1.1 uses a quadrupole mass filter to separate the fragments after photofragmentation in the ion trap. There are two main disadvantages of this configuration: relatively low mass resolution and the need to scan the quadrupole in order to observe the parent ion and all fragments. Increasing the mass resolution would be beneficial, for example, because under "gentle" source conditions dimers and other oligomers of the parent molecules can be observed (see Chapter 4), and they can be distinguished using the spacing of the isotope pattern. One solution would be to combine the CIS instrument with a TOFMS. It is technically challenging to combine a TOFMS with a linear ion trap, due to a broad spatial distribution of the ions exiting the trap. For this reason we changed the trap geometry and introduced a planar RF ion trap with an orthogonal ejection.[284]

The design of a planar trap was originally suggested by Wester[285] and further developed by Lorenz and co-workers.[286,287] Two sets of 16 electrodes form two planes, separated by a 0.5 cm gap (Figure 2-9). They are manufactured on printed circuit boards (PCBs) with gold coating. The RF electrodes are surrounded by 4 DC plates for ion confinement (Figure 2-9 (b)). The PCBs are mounted between two copper plates, and are in thermal contact with a copper frame, which is connected to the second stage of a helium cryostat (Sumimoto, RDK-408E2, Darmstadt, Germany). Ions are extracted through slots cut between the RF electrodes on the upper PCB into the TOFMS.[288] The entire cold trap is enclosed in an aluminum heat shield, which is attached to the first stage of the cold head and kept at ~40 K.

The design of the IMS separation stage requires that the ions begin their motion at the entrance of the drift tube at the same time, as a packet, imposing that the instrument is run in a pulsed mode. The frequency of the IR laser pulses is 10 Hz, which determines the optimal length of the cycle of the machine – 100 ms. Several milliseconds before the ion packet enters the cold trap, a pulsed Parker Series 9 valve introduces there a He/H_2 gas mixture. To prevent H_2 from condensing on the walls of the chamber and the electrodes, we keep the temperature of the trap at 13 K through the use of a Lake Shore 350 PID controller and a cartridge heater. The RF voltages of opposite phase are applied to alternating electrodes on each PCB. The ions enter the trap and cool down in collisions with the gas mixture, and as their temperature decreases, H_2 clusters including analytes are formed. The

voltage on the entrance DC plate is kept low while the ions are arriving into the trap and then gradually increased as the ions cool down in order to confine them in the center of the trap.
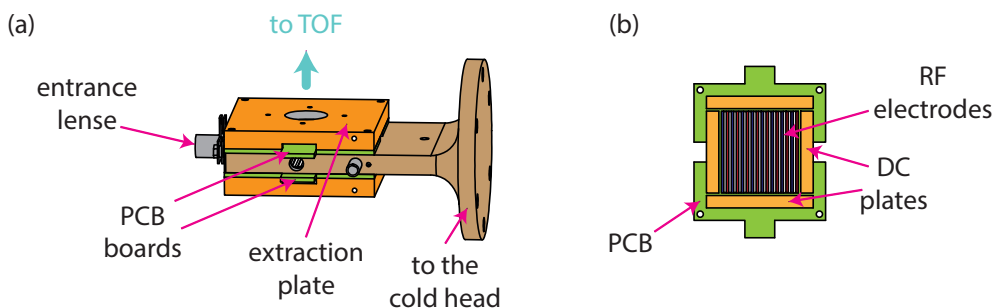

Figure 2-9. (a) Schematic of the flat trap design. (b) Schematic of the PCB board used in the flat trap design.

After ~50 ms, the ions are equilibrated and form a cloud of ions clustered with hydrogen in the center of the trap, while most of the gas mixture is pumped out. At this moment the RF voltage on the electrodes is switched off, and the clusters are extracted through the upper PCB (see Figure 2-9 (a)) towards the TOF. This is achieved by applying ~3.5 kV to the bottom RF electrodes, ~2 kV to the top RF electrodes and ~0.8-0.5 kV to the top extraction plate. Upon exiting the trap, the clusters can dissociate in collisions with the residual gas, producing bare parent ions. An example of cluster size distribution recorded in the TOF is shown in Figure 2-10.

The clusters are interrogated with a laser beam ~2 µs after their extraction from the trap, directly above the heat shield. A 10 ns IR laser pulse is produced by an OPO-OPA (Laser Vision), pumped with a Nd:YAG fundamental (Innolas). The line width of the laser is ~3 cm^{-1}, and for small systems it is the limiting factor in the resolution of the obtained spectra. The spectra presented in Chapter 3 of this thesis have been obtained before the latest upgrade of the instrument, which has significantly decreased the temperature of the ions. In this case, the width of the observed bands was limited by the temperature of the ions, which was ~ 30 K.
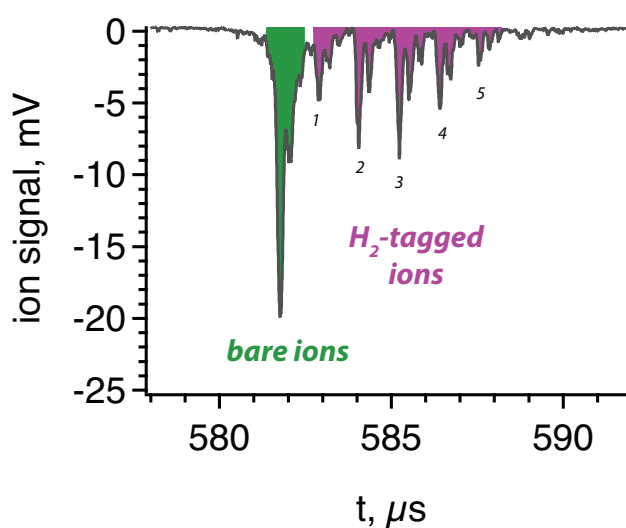
Figure 2-10. A typical TOF mass-spectrum featuring the bare analyte ion shown in green and a distribution of H_2-tagged species in pink. An infrared spectrum is recorded as depletion of the number of tagged molecules as a function of the laser wavelength.

A general description of H_2-tagging spectroscopy is provided in section 2.1.11. Practically, the laser excitation reduces the mass of a cluster by boiling off one or more H_2 molecules. This mass change is detected in a 1 m long reflectron TOFMS (Jordan TOF Inc.). The data acquisition is controlled by in-house software (LabView). The signal from the MCP is recorded on a LeCroy oscilloscope, which is then read by a PC. The program then averages a pre-set number of data sets and integrates the signal to find peak areas for the clusters and the parent ions (see Figure 2-10). This information is used to reconstruct the IR absorption spectrum of the species.

2.1.11 Principles of tagging spectroscopy

We acquire IR spectra in the hybrid IMS-CIS instrument using H_2 pre-dissociation, which was originally proposed by Okumura[289,290] and further developed by Johnson.[209,291] Depending on the temperature of the trap and other technical limitations, a number of chemically inert gases can be used as messengers.[225,292-295] The two main requirements for the "tag" molecule are:

- The interaction of the messenger with the molecular ion should be weak and not perturb the structure of the molecule.

- The sticking probability of the messenger should be high enough to observe sufficient number of clusters.

While He is the optimal choice to fulfill the first requirement, the temperature of the trap has to be close to 4 K to produce a sufficient amount of clusters. On the other hand, the inert gases with high polarizability such as Ar or Ne readily attach to the charges of the molecular ion,[295-297] but they significantly perturb its vibrational spectrum.[293,298] Hydrogen appears to be a compromise solution that satisfies both criteria.

The effective binding energy of H_2 to glycine was found to be 840 cm^{-1}, to a model dipeptide, 490 cm^{-1}, and to a tripeptide, 370 cm^{-1}. The difference between these values is determined by the accessibility of the charge sites rather than the ion size.[209,291] These binding energies allow attachment of 1-14 messenger molecules in our cold trap, though the shape of the cluster size distribution highly depends on the settings of the instrument: the gas pulse timing, ion trap pole bias, etc.

The previous studies using H_2 as a messenger suggested no significant changes in the vibrational spectra of the ions upon tag attachment.[209,291,294] However, a thorough comparison between He and H_2 shows that H_2 can perturb the gas-phase structure of a small amino acid such as glycine.[299] One H_2 molecule produces a red shift in the frequencies of the NH_3 bands, which is consistent with H_2 binding to the charge site. Furthermore, attachment of three tag molecules induces a conformational change in glycine by breaking a weak hydrogen bond and thus forming a second conformer, which becomes dominant in clusters with six or more tag molecules. With that in mind, in this work we perform the measurements using only one messenger molecule attached to the parent ion. In our case, however, further comparison between double-resonance spectra of large molecular ions and their spectra obtained *via* tagging spectroscopy shows no significant differences for the clusters of up to 10 H_2.

In principle, a spectrum is recorded as depletion of the number of clusters. At the same time, as the tags are removed, the number of bare parent ions is increased, so the normalization is performed as follows:

$$spectrum = \frac{sum\ of\ signal\ from\ all\ clusters}{sum\ of\ signal\ from\ all\ clusters + parent\ ion\ signal}. \tag{2.24}$$

The normalization accounts for possible fluctuations in the ion current during the time of one laser scan and improves the signal-to-noise ratio.

2.2 Nuclear magnetic resonance

The nuclear magnetic resonance spectra are acquired at room temperature on a Bruker Avance III HD instrument operating at 600 MHz for ^{1}H. Two solvents are used: DMSO and a H_2O:CD_3OH 50:50 mixture. The former is aprotic and does not interfere with the signal from the sample, which simplifies spectral assignments. The latter solvent reproduces the conditions used in the electrospray process. The use of fully deuterated solvents is not possible due to fast H/D exchange, but protons on the solvent molecules that do not undergo H/D exchange are substituted by deuterium to decrease the interference with the sample signal.

We first acquire TOCSY and COSY 2D ^{1}H NMR spectra in DMSO in order to fully assign the peaks of the main conformer. This helps interpretation of analogous data obtained in a H_2O:CD_3OH mixture, because although water suppression was performed with a pre-saturation sequence, the signal from solvent protons complicates the spectral assignment. The ROESY experiment (mixing time 400 ms) identifies the protons that are located at distances of 2-5 Å and allows one to deduce the structure of the peptide and the conformation of the proline residues. The chemical shifts are compared to those observed for the unstructured protein regions, and the relative abundances of the conformers are measured by integrating the peak areas. The data analysis was performed using Mestrelab MNova NMR software.[300]

2.3 Conformational search

Infrared spectra are spectroscopic signatures or, in other words, fingerprints of biomolecular conformations. However, to interpret the information contained in them, one needs to perform quantum chemical simulations and compare the computed IR spectra with those experimentally measured. Other information, such as CCSs and UV absorption, can be used to guide and verify the simulations.

2.3.1 Creating the initial pool of structures using force field

We generate an initial pool of calculated structures using basin-hopping, as implemented in the Tinker package[301] with OPLS-AA force field.[302] The energy cutoff was set to 50 kcal/mol, and the number of torsional modes was set to 20. Increasing either of these parameters does not lead to a substantial increase in the number of unique structures generated. The initial pool of structures was subjected to a conformational clustering

53

procedure from the GROMACS program package,[303,304] which is based on Cartesian RMSD value (cut-off 0.75 Å). Only the lowest-energy representative of each cluster was kept.

2.3.2   Collisional cross section calculation

The seemingly trivial issue of CCS calculation has attracted much attention in the recent years. There are several methods available ranging from very fast but imprecise to the most accurate, for which the uncertainty is on the order of the experimental error.

- The Projection Approximation (PA)[305] is historically the first and the most intuitive way of estimating the CCS. It represents the CCS as a rotationally averaged area of the shadow of the object. While computationally efficient, it usually underestimates CCSs for polyatomic species[306], and it is typically used for small convex molecules consisting of up to 100 atoms. PA is implemented in several software packages, including MOBCAL[307] and Sigma.[308] For larger systems, appropriate scaling factors can be found using calibration against more precise methods.[309]

- Further development of the PA method led to the projection superposition approximation (PSA),[310,311] which takes into account (i) the cooperative size effect by superposition of diffuse atomic contributions and (ii) the cavities using a shape factor.[312] PSA can be used both for He and N_2 drift gases.[313]

- Exact hard sphere scattering (EHSS)[314] is used for molecules with concave surfaces. In this method the atoms are considered as hard spheres and the trajectories of the gas molecules resulting from elastic collisions with the analyte are calculated explicitly. Due to multiple collisions of a buffer gas molecule with the ion, the CCS can increase drastically compared to the PA.[315]

- The Trajectory Method (TM)[305] takes into account the long-range interactions between the analyte and the gas molecules. It is the most direct and precise commonly used way of describing the CCS but the most computationally expensive.[306] Usually a 6-12 Lennard-Jones potential is employed along with charge-induced dipole interaction terms, which can be parameterized for different atoms and drift gases. Originally the parameters were obtained from carbon cluster measurements in helium, while improved sets of parameters

for the atoms different from C and for nitrogen as a drift gas were obtained later.[316]

- Further improvements of the model include non-elastic non-specular collisions.[317]

In this work the CCSs first were calculated using the PA method[305] implemented in Sigma.[308] Knowing that this method underestimates the collisional cross-sections,[89,310] we computed the CCSs of the selected pool of ~4,500 structures with the trajectory method (TM[318] in MOBCAL[307], the partial charges obtained from the force-field parameterization), and used linear regression to find a correspondence between CCSs computed by PA and TM for our system (see Chapter 3). In later stages we calculated the CCSs using the TM method with partial charges extracted from the DFT calculations.

2.3.3 DFT optimization and vibrational spectra simulation

All DFT simulations in this work were performed in the all-electron program package FHI-aims based on numeric atom-centered orbital basis sets.[319] Initially the PBE functional[320] with many-body dispersion correction (MBD)[321,322] was used. At first, relaxations were performed with the "light" species defaults, while for refinement a more accurate "tight" basis was employed. For further improvement of the description of the system we used the PBE0 functional.[323] These functionals, and the order of increasing accuracy, were proven successful in previous studies.[89,324-326]

The vibrational spectra are first simulated using the PBE functional and then for several selected structures re-computed using more accurate PBE0 functional. To correct for the anharmonicity of the vibrations in the NH stretch region we used scaling factors of 0.948 for PBE0 and 0.978 for PBE.

Free energies were estimated, based on the vibrational frequencies calculated with the PBE functional, at 0 K, 10 K and 300 K using harmonic oscillator and rigid rotor approximations.

Chapter 3.
Conformations of prolyl-peptide bonds in the bradykinin 1-5 fragment

As the first step in studying kinetically trapped conformations of intrinsically disordered peptides in the gas phase, we address the structural preferences of the doubly protonated bradykinin 1-5 fragment, RPPGF. We combine ion mobility with cryogenic ion spectroscopy to study the peptide in the gas phase, perform extensive quantum-chemical simulations to interpret and complete the experimental results, and employ nuclear magnetic resonance to get an insight into the structure in solution and compare it to that in the gas-phase. The results presented in this chapter were partially published previously.[327]

3.1 Introduction

Information about the structure and dynamics of proteins and peptides is crucial for understanding their physiological function and hence essential for diagnostics and drug design.[59,328] X-ray crystallography and NMR spectroscopy can often successfully determine structure of biomolecules when they adopt well-defined secondary structures.[28] However, in the case of intrinsically disordered peptides (IDPs), NMR spectra often represent an average over an ensemble of different conformations. Moreover, crystallization for X-ray measurements is often not possible due to the highly dynamic character of the disordered states.[39,51,52] Many questions thus remain open regarding the nature of intrinsically disordered peptides.

A powerful way to obtain an insight into the nature of intrinsically disordered peptides is to study them in the gas phase in the absence of solvent.[66,89,96,145,146,211,329-332] The expected distribution of structures produced by electrospray is a mixture of low-energy gas-phase conformers and higher energy structures that are kinetically trapped, with barriers on the potential-energy surface (PES) that prevent isomerization.[74,83,136,195,333-336] These latter structures are particularly important, as they retain information on the conformation of the molecule in solution (see Chapter 1 for more details). In the case of proline *cis-trans* isomerization, the energy barrier can be high enough (20-50 kJ/mol)[74] to expect solution-phase isomers to be preserved as metastable species. Detailed understanding of these kinetically trapped conformations would allow one to connect information obtained from solution and gas phase techniques.

Low-energy gas-phase conformers, on the other hand, may contain structural features that are produced upon change of environmental conditions as, for example, during membrane insertion. Membranes are generally seen as structure-inducing media, and a question was recently raised if molecular disorder is possible there.[63] Detailed conformation studies in the gas phase might shed light on this question, since vacuum can be understood as an extreme case of an aprotic non-polar solvent lacking intermolecular interactions.

Although much insight can be gained from IMS-MS, it provides only an average CCS, which is not unambiguous for a conformational family in the gas phase.[337] A single IMS peak may contain several distinct conformational families with close CCSs. To "zoom in" on the details of molecular structure, spectroscopic techniques can be combined with mass spectrometric ones.[338] The most informative is the combination of spectroscopy with IMS, which is realized in the hybrid IMS-CIS instrument (section 2.1.10). We also employ double-resonance IR-UV spectroscopy to provide complementary to IMS, conformer-specific information (see section 2.1.2).

An infrared spectrum can be seen as a vibrational fingerprint of a molecule's conformation, but it cannot be interpreted without high-level quantum-chemical simulations, which, in turn, have to be verified by comparison with the experimental data. The latter is particularly true for kinetically trapped species, which by definition lie high in energy and thus cannot be identified using energy minimization. At the same time, the results of gas phase experiments have to be compared with solution phase studies in order to examine the potential correlation between the structure of kinetically trapped species

and of those found in solution. To that end, we also perform experiments using NMR spectroscopy. When all techniques listed above are merged, they produce a general overview of the conformational space of a molecule with a high level of detail in solution and in vacuum. Thus each method benefits from its conjunction with other approaches (Figure 3-1).
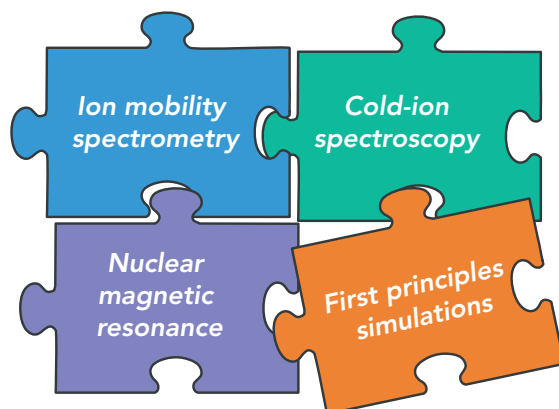

Figure 3-1. List of techniques used in this study. The jigsaw represents the productive use of the methods in combination: IMS helps separating conformational families to simplify vibrational spectra and to guide the conformational search; theoretical simulations allow interpretation of the vibrational spectra; spectroscopic tools provide structural information beyond CCSs and verify the results of the simulations; NMR provides the solution phase structural information for comparison.

We employ this combination of experimental and theoretical methods to study the 1-5 fragment of BK (Figure 3-2), which is metabolically stable[198] and was shown to inhibit the enzymatic activity of thrombin.[199-201] Experimental constraints are used to guide the first-principles structural search and allow us to identify not only the thermodynamically stable low-energy conformers, but also those that are kinetically trapped. We show that *cis-trans* isomerization of BK[1-5] plays a key role in forming distinct conformational families and that the kinetically trapped structures have a direct relationship to those observed by NMR in solution.
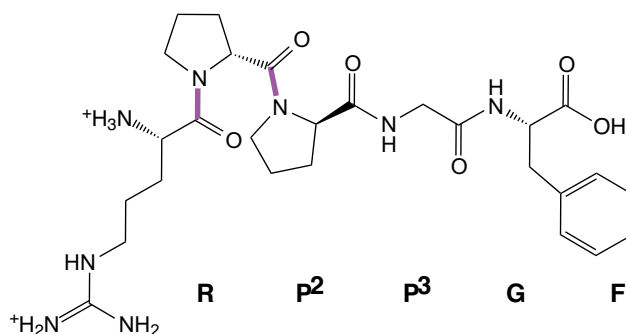
Figure 3-2. Chemical structure of doubly protonated bradykinin residues 1 to 5 (BK[1-5]$^{2+}$). The bonds around which the *cis-trans* isomerization can occur are highlighted.

3.2 Bradykinin 1-5 in the gas phase: experimental study

3.2.1 *Ion mobility distribution of bradykinin 1-5*

The ion-mobility drift-time distribution of BK[1-5]$^{2+}$ has two well separated peaks, as shown in Figure 3-3. The first peak has an unresolved shoulder and can be represented as a sum of two Gaussians, one centered at 167 Å^2 and the other at 170 Å^2. The second peak is centered at 178 Å^2, and its width suggests that it may also consist of more than one conformer. As shown in Figure 3-4, when we select either of the two peaks after they drift through the first part of the drift tube and collisionally activated them, the same arrival-time distribution is produced for both peaks in the second segment of the drift tube. This annealed distribution consists mainly of peak I with a shoulder around 176 Å^2. We conclude that the structures contained within peak II are largely kinetically trapped, while peak I represents stable gas-phase structures.
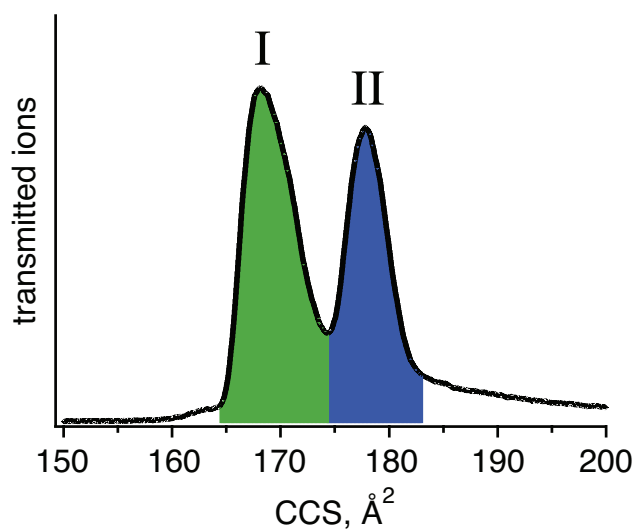
Figure 3-3. Collisional cross-section distribution of BK[1-5]$^{2+}$ (RPPGF). The extended conformers (peak II) are kinetically trapped and interconvert to more compact, stable, gas-phase structures (peak I) upon collisional activation.
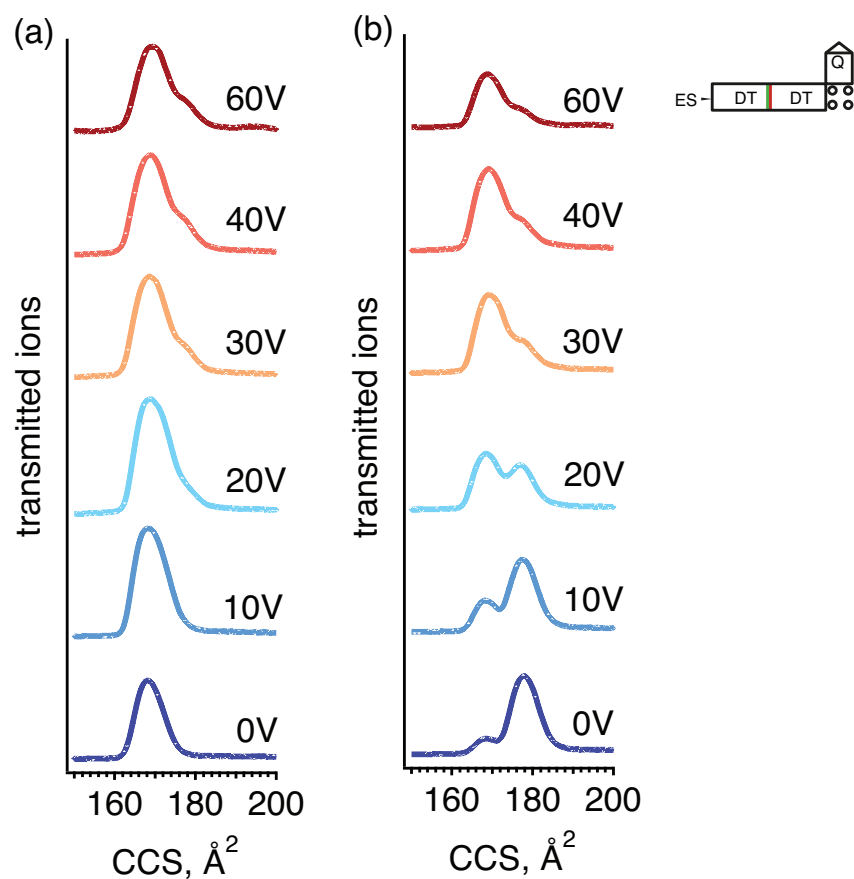

Figure 3-4. Collisional activation of BK[1-5]$^{2+}$ pre-selecting (a) peak I and (b) peak II in the drift-time distribution of Figure 3-3.

61

3.2.2   Ion mobility distribution with higher resolution

The IM distribution shown in Figure 3-3 and all other experiments, except for the ones presented in this section, were performed using the first generation of the IMS-CIS instrument, in which the resolution of the drift tube has not yet reached the values reported by Clemmer for the same instrument (R~100).[281]   Later we have measured the IM distribution of BK[1-5]$^{2+}$ with maximum resolution achievable in our instrument, which reveals several new details (Figure 3-5). Peak I indeed consists of two peaks with close CCSs, 167 Å^2 and 168 Å^2, while peak II spans from 175 Å^2 to 182 Å^2. A closer study of this distribution shows that the broad peak II might contain solvent adducts or very extended high-energy metastable conformations that are preserved in the drift tube and dissociate/interconvert at the end of it upon extraction.
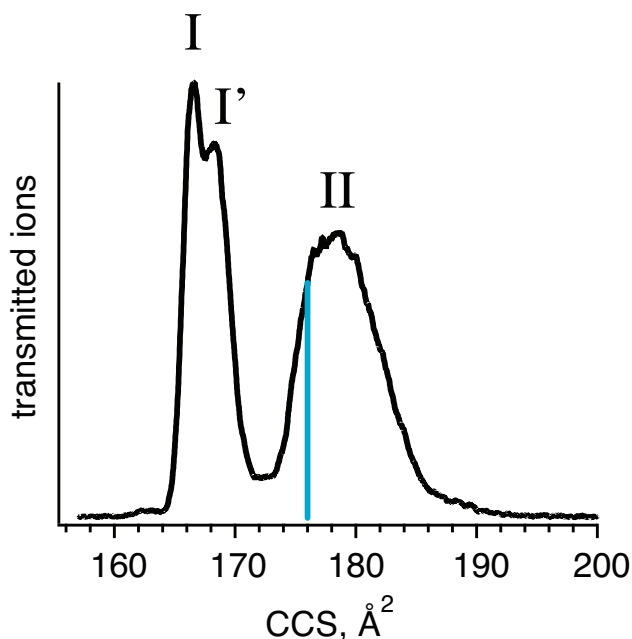
Figure 3-5. Ion mobility distribution of BK[1-5]$^{2+}$ obtained using "gentle" source conditions. The blue line shows the CCS of the bare ion with extended conformation (see text for details).

There are two ways to check if a certain peak or a shoulder at high CCS is formed by clusters with the solvent molecules rather than a bare ion:

- If solvent adducts are strongly bound, the IMS distribution recorded on the masses of the adducts should share the high CCS peaks with the IMS distribution of the bare ion. This is the case if part of the solvent adducts dissociates at the end of the drift tube, causing these peaks to appear in the

drift-time distribution of the bare ion. An example of this situation is presented in Chapter 4.

- If the solvent adducts are weakly bound, all of them might dissociate at the end of the drift tube. In this case we would not observe any of them in the TOF mass spectrum, but we can get an indication that the high CCS species *might* be adducts with solvent molecules by using the selection gate.

Normally, the time, during which the ions drift through the entire length of the drift tube, is proportional to the time they require to get to the selection gate:

$$\frac{t_{full\ drift}}{t_{selection\ gate}} = \frac{l_{full\ drift\ tube}}{l_{first\ section}} = const, \qquad (3.1)$$

since the velocity of a given ion is determined only by the electric field. For example, for triply protonated BK we measure:

$$t_{full\ drift} \approx t_{selection\ gate} * (2.33 \pm 0.04). \qquad (3.2)$$

Applying the selection pulse slightly activates the ions, and if the clusters are weakly bound, they dissociate and change their mobility, causing the above relation not to hold.

Consider first the compact IM peaks I and I` (Figure 3-5): for them the relation between the total drift time and the time of the selection holds, as shown by the red line in Figure 3-6. It also holds for the compact shoulder of peak II. However, the extended structures interconvert to the more compact ones upon applying the selection pulse, and the relation between the total drift time and the selection time changes (blue curve in the Figure 3-6). We determine that all extended structures require the same time (4.9 ms) to drift through the second part of the instrument, which means that they all contract to the same CCS upon selection. Converting the drift time to the CCS, we see that any adduct produces a structure with CCS equal to 176 Å^2 (shown with a blue line in Figure 3-5).

This analysis suggests that we might be observing solvent adducts (we use 50:50 methanol:water mixture in this experiment) along with metastable extended structures, the latter of which we are interested in. In either case, upon selection in the middle of the DT these species convert to a conformational family with CCS equal to 176 Å^2, which is eventually loaded to the cryogenic trap for further spectroscopic analysis.
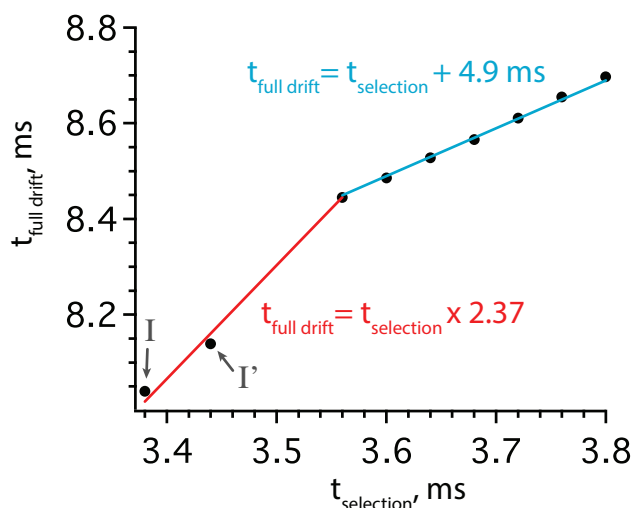
Figure 3-6. The total drift time of selected slices of DT distribution of BK[1-5]$^{2+}$ as a function of their selection time. Normally, the two times are directly proportional if no activation is performed (red curve). If upon selection the weakly bound solvent adducts dissociate, their collisional cross-section shrinks and the total drift time follows the equation shown in blue.

3.2.3   Vibrational spectra of mobility-selected ions

While ion-mobility spectrometry provides a spatially averaged CCS of a molecule that reflects its overall shape, vibrational spectroscopy produces a distinct fingerprint, which reflects the molecule's precise structure. We thus select each peak in the ion-mobility distribution and obtain its IR spectrum *via* H_2-tagging spectroscopy (Figure 3-7). We also record their vibrational spectra after annealing *via* collisional activation (50 V across the activation region), and observe that they are practically identical and contain the same bands as the vibrational spectrum of the compact structures (Figure 3-8). This confirms that after the collisional activation we reach the same quasi-equilibrium gas-phase distribution independent of the starting conformation. This distribution mainly consists of the same conformers as those under the more compact peak I in the ion mobility drift-time distribution.

The spectrum of peak II contains traces of the IR transitions of the annealed distribution due to collisional activation at the end of the drift tube, which is difficult to avoid completely. The most intense bands are found at 3540-3565 cm^{-1} as well as at 3440-3465 cm^{-1} and 3500 cm^{-1}. To obtain a vibrational spectrum of the species at 176 Å^2 without contamination we subtract the spectrum of the annealed distribution (multiplied by an appropriate coefficient) from the raw spectrum of peak II. It results in the blue spectrum, shown in Figure 3-8, which we use for further analysis.
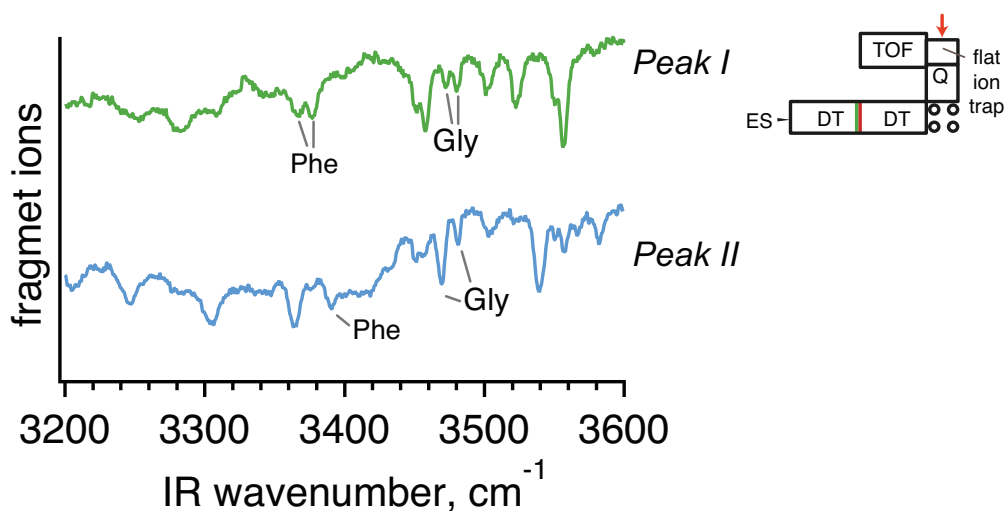
Figure 3-7. Vibrational spectra of mobility-selected BK[1-5]$^{2+}$. Phenylalanine and Glycine NH stretches are assigned by measuring the spectra of ^{15}N-labeled peptides.
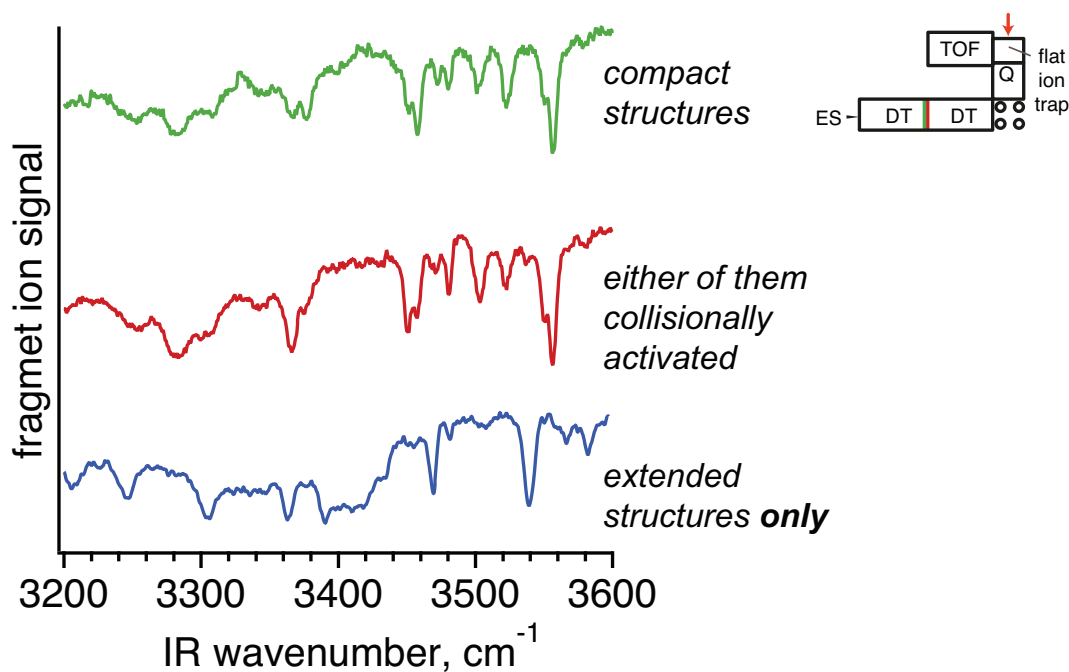
Figure 3-8. Vibrational spectra of mobility-selected BK[1-5]$^{2+}$. Annealed distribution is obtained by collisional activation and its spectrum is independent of the initially mobility-selected peak.

To assign the peaks in the IR spectra to specific vibrational modes, we had two isotopically labeled peptides synthetized by replacing the amide nitrogen of either phenylalanine or glycine with ^{15}N. This substitution typically shifts the labeled NH band by ~8 cm^{-1} to the red[213] and allows us to identify the bands that correspond to phenylalanine

65

and glycine NH stretches, which are labeled in Figure 3-7. These vibrational spectra represent a mixture of several conformations, as we discuss in more detail in the next Section, therefore in some cases we observe two bands per oscillator.

3.2.4 *IR-UV double resonance spectroscopy*

One of the most powerful ways to simplify congested spectra of large molecules is to use double-resonance spectroscopy in a cryogenic ion trap.[211] We first record a UV photofragmentation spectrum, shown in Figure 3-9 in our cold, octopole ion trap and then performed IR-UV double resonance spectroscopy parking the UV laser on all major peaks in this spectrum. This way we record vibrational spectroscopic signatures of three distinct conformers, shown in Figure 3-10, and label them A, B and C. Isotopic labeling allows us to identify the NH stretches of the phenylalanine and glycine residues. Note that the IR spectrum of conformer B is very close to that of conformer A, with the main difference being the frequencies of the phenylalanine NH stretches, which are separated by 27 cm^{-1}. This suggests that these two conformers differ in the rotation of phenylalanine side chain.[91,213]
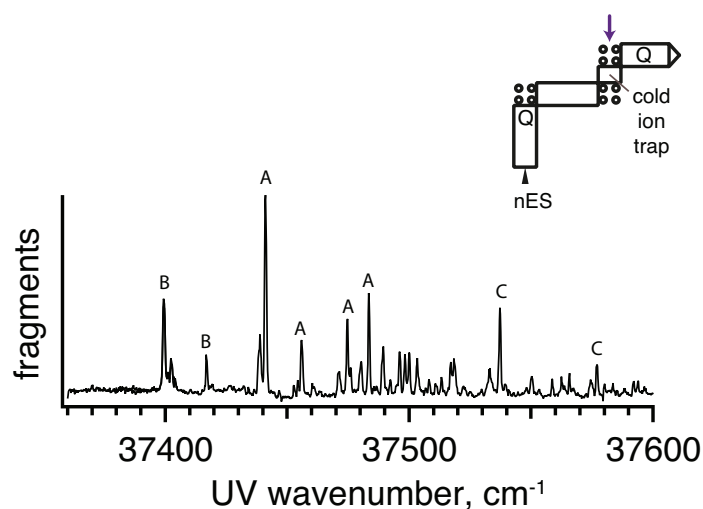

Figure 3-9. Electronic photofragmentation spectrum of BK[1-5]$^{2+}$. The letters label the peaks corresponding to different conformations.
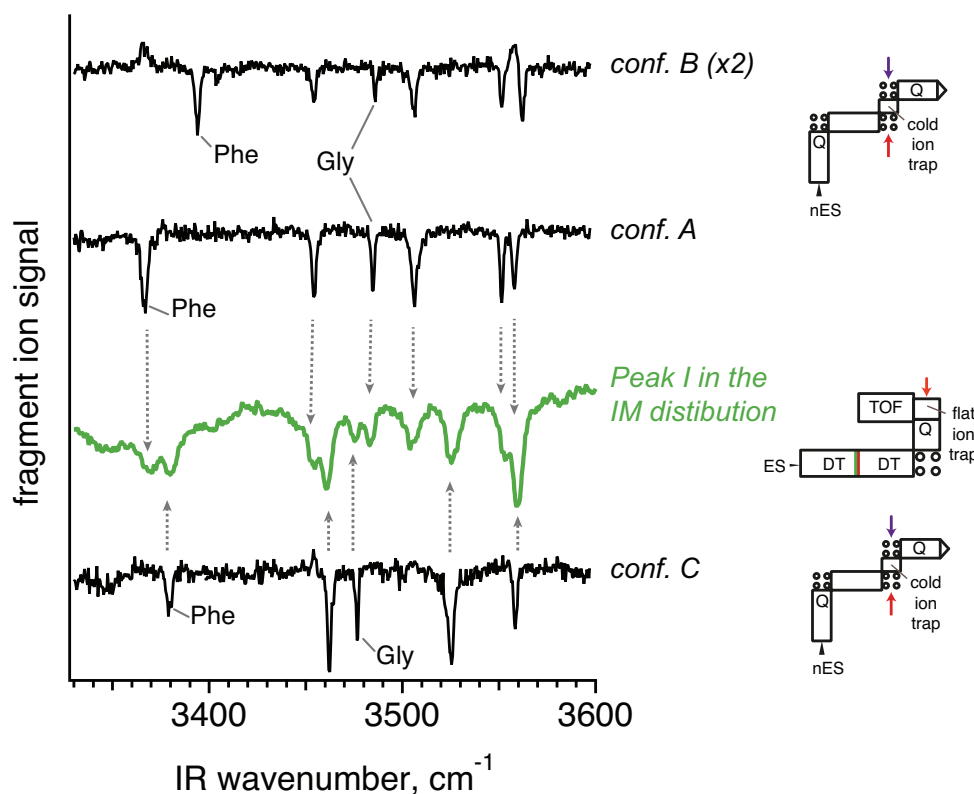
Figure 3-10. Comparison of vibrational spectra of 3 conformers (A, B and C) of BK[1-5]$^{2+}$ obtained in the CIS machine by IR-UV double-resonance together with a vibrational spectrum (in green) of the mobility-selected compact structures recorded in the IMS-CIS machine by H_2-tagging. The low-CCS peak in the arrival-time distribution, accordingly, consists of a mixture of conformers A, C and, possibly, B. The labels show the NH-stretches that were assigned by measuring the spectra of ^{15}N-labeled peptide.

Additional geometrical constraints can be obtained using vibrational spectra of the electronically excited molecules in the S_1 state.[226,227] As electronic excitation is local, only the vibrations in close proximity to the chromophore shift in the IR spectrum. We confirm this statement for the phenylalanine NH-stretches, previously assigned using isotopic labeling, for all three conformers (Figure 3-11). We also observe a small, but reproducible shift in the position of the band at 3558 cm^{-1} for conformer A, which suggests that this band belongs to the carboxylic OH stretch, rather than to the arginine side chain.

The spectra of Figure 3-10 clearly demonstrate that peak I identified by ion mobility consists of conformers A and C. At the same time the spectroscopic differences between spectra of A and B are so small that we cannot exclude presence of small quantities of conformer B. In the CIS instrument that we currently use, ions undergo collisional activation at the end of the ion funnel. This explains why we observe a set of structures that largely resembles the annealed distribution obtained in our IMS-CIS machine, i.e. mostly the compact structures. Despite the recent finding that H_2-tagging can alter the

67

conformation of small biological molecules,[299] in this case we observe that the spectroscopic features are not detectably shifted when the peptide is tagged with one H_2 molecule.
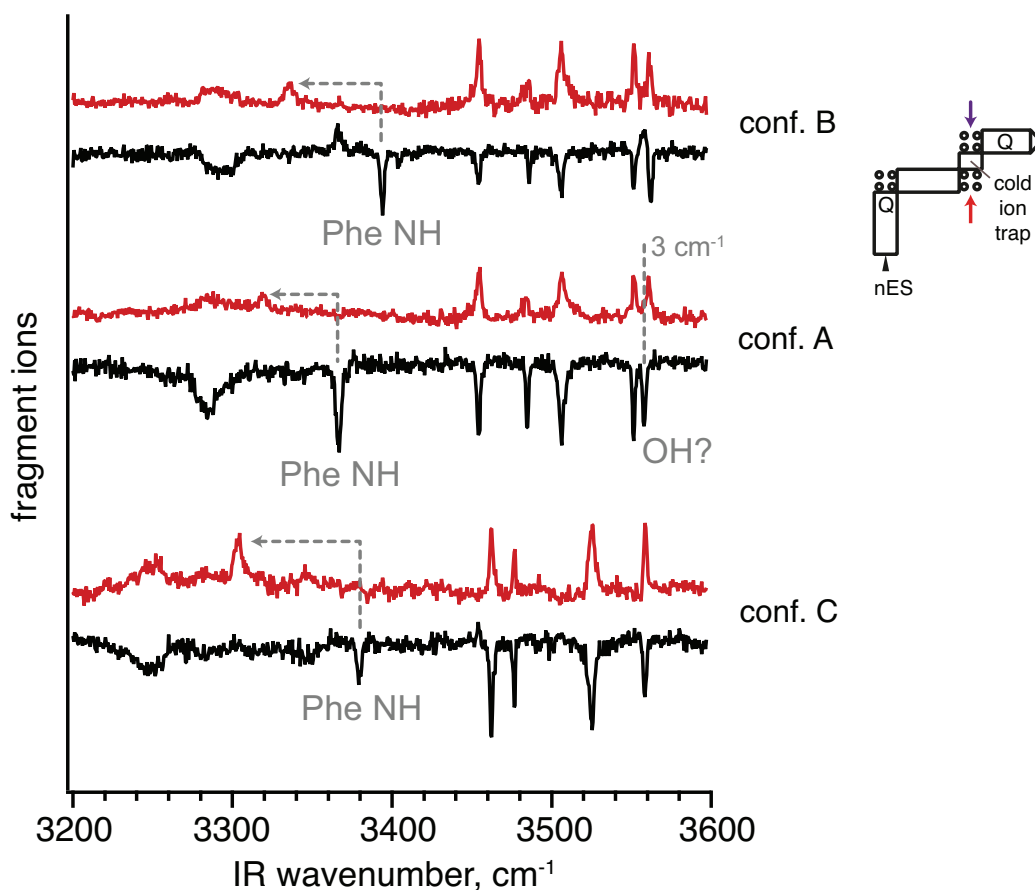
Figure 3-11. Vibrational spectra of three conformers of BK[1-5]$^{2+}$ in the electronically excited state (red), compared to their vibrational spectra in the ground state (black). The shifting bands are labeled.

Having the vibrational bands in the spectrum of peak I assigned to particular conformations, we can observe the changes that happen within the conformational population of peak I upon collisional activation. Figure 3-12 shows that the annealed distribution consists of the same conformers, A and C, as peak I, but their relative intensities differ slightly: upon activation conformer C converts to conformer A.
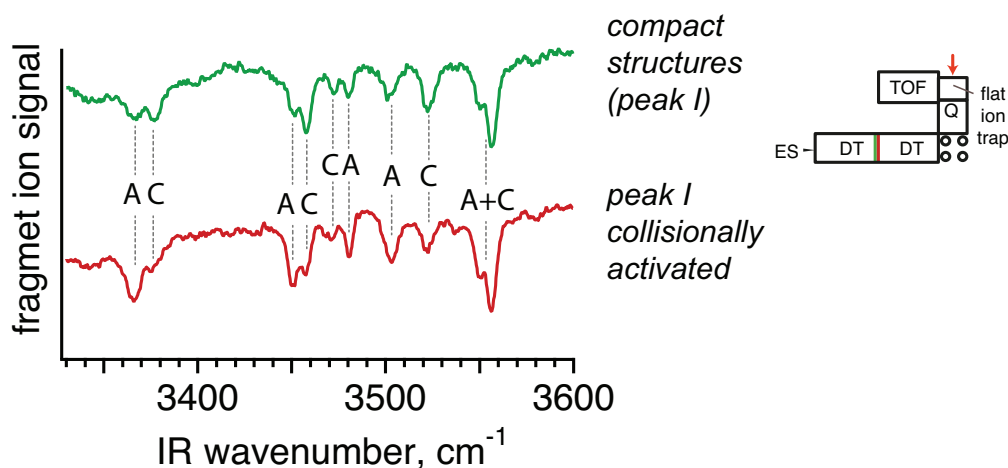
Figure 3-12. The annealed distribution consists of the same conformers as peak I in the drift-time distribution, but their relative abundances change, as can be seen from different relative intensities of the bands: conformer C converts to conformer A.

To summarize the experimental results, the following structures are observed in the gas phase:

- Compact conformations with a CCS in the range of 166 to 168 $Å^2$. Analysis of their spectral signatures suggests the presence of conformers A, B and C. B only differs from A in the phenylalanine ring orientation. A third conformer, denoted C, is structurally different from A and B.

- Extended structures with a CCS around 176 $Å^2$. They are kinetically trapped and upon collisional activation collapse to the more stable compact conformations.

Previous studies of the full BK sequence suggest that families I and II differ by prolyl-peptide bond isomers.[70,173]

3.3 First-principles simulations

To determine the structures of the conformers that give rise to the CCS values and infrared spectra, we have to compare the experimental results with theory. First we must search the massive conformational space for BK[1-5]$^{2+}$ (discretization of the 13 single bonds in 60 degrees steps and of the 4 peptide bonds to *cis* and *trans* states results in roughly 2×10^{11} structures to evaluate). We tackle this sampling problem by a two-step approach: (i) a force field-based global screening (basin hopping with Tinker and the OPLS-AA force field), and (ii) subsequent density-functional theory calculations. In order to select the conformers to consider at the higher, first-principles level, we employ experimental constraints, comparing the calculated CCS values and vibrational signatures

69

to their experimental counterparts. The use of experimental constraints is not only a concession to the large conformational search space. Only one of the observed conformers can be the global minimum of the potential-energy surface, while the rest, especially the kinetically trapped species, cannot be identified based on the energy criterion alone.

The initial force field search yields 212,754 conformations that were sorted into 67,546 clusters, the lowest energy representative of which was considered in the following. We first assumed, and later confirmed, that kinetic trapping is a result of the *cis/trans* isomerization of the two prolyl-peptide bonds present in BK[1-5]$^{2+}$. All structures were categorized into four groups, *trans-trans* (TT), *trans-cis* (TC), *cis-trans* (CT), and *cis-cis* (CC), according to their prolyl-peptide bond conformations. We have carried out the following analysis separately for each category so that we are sure to consider structures that are high in energy, such as TT, but that may still be observed experimentally due to kinetic trapping. We selected structures out of these individual sets based on their relative potential energy and calculated CCSs. In order to put an emphasis on structures with CCSs that match the experiment, we applied an energy threshold of 22 kcal/mol (relative to the lowest energy structure) to structures with CCS values between 160.5 Å^2 and 178 Å^2 (computed with the PA method). For structures with a lower or higher CCS, we applied a relative energy threshold of 18 kcal/mol, as illustrated in Figure 3-13. By this procedure, we selected 4,515 structures that were subjected to geometry relaxations using the PBE functional with many-body dispersion correction and "light" computational settings.

We then computed the CCSs for all 4,515 DFT-optimized structures using the PA and TM methods in order to access the applicability of PA as a faster but less accurate procedure (Figure 3-14). Recovering the TM CCSs using the PA values and a linear "calibration" produces an acceptable error of ~1.5 %. Using this calibration we can also estimate retrospectively the "true" CCS limits, which we used to select the structures for the DFT optimization. They turn out to be 162-179 Å^2, as shown with the red line in Figure 3-13.
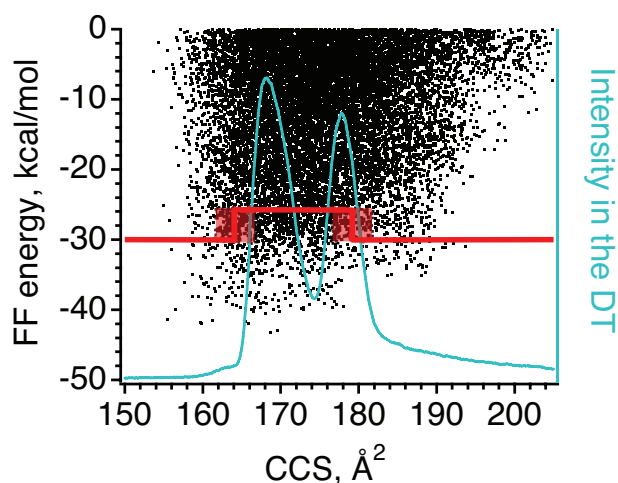
Figure 3-13. The energy of structures unidentified by force field conformational search as a function of their CCS (black dots). The energy cutoff for the DFT optimization depending on the CCS is represented by a red line. The uncertainty coming from using PA for the CCS calculation is ~1.5% and it is shown with the semi-transparent red borders.
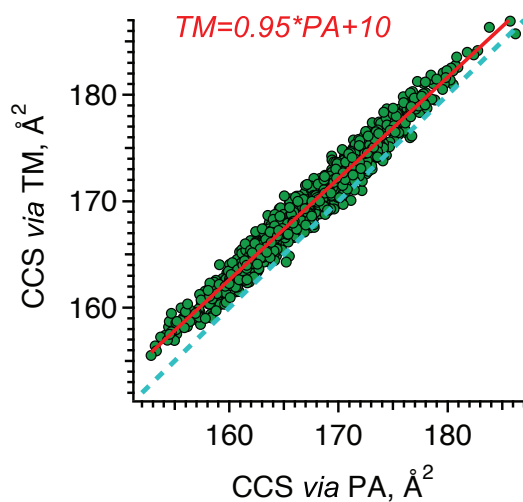
Figure 3-14. Comparison between CCSs calculated using TM and PA methods with the linear fit. The dashed line of equal CCSs guides the eye.

It is instructive to compare the energies and the CCSs of the computed structures before and after the DFT optimization (Figure 3-15).

Figure 3-15. (a) The energies of all CT structures after the DFT optimization with the "light" settings as a function of the FF energy. (b) Comparison between CCSs calculated after force field and DFT optimization of the same geometries. The dashed line of equal CCSs guides the eye.

The prediction of the lowest-energy structural type changes when going from the FF to the DFT: while the former assigns a TC structure to be the lowest-energy, the latter predicts a CT structure to be a global minimum on the PES. In general, the FF energy value hardly predicts the energy after the DFT optimization (Figure 3-15 (a)). There are two main reasons for this:

- First, the geometry of the molecule changes substantially upon DFT optimization, as can be seen from Figure 3-15(b). The general trend is the decrease in the energy upon re-optimization with DFT, probably, due to several FF structures converging to one minimum on the DFT PES. This explanation is supported by the fact that clustering of the structures after the DFT optimization produces 33 % less clusters, i.e. distinct conformations, than before the DFT. The geometries also become on average slightly more compact.

- Secondly, the algorithm of the energy estimation is fundamentally different in the DFT and FF simulations.

It should be noted that since the vibrational spectra require *a priori* a high level of theory, we did not optimize the FF to achieve the highest precision of energy estimation but rather used the FF conformational search to generate the initial pool of structures for further DFT analysis.

3.3.1 Compact structures: Low-energy minima

To find candidates for the equilibrated stable structures of peak I, we rely mainly on the potential energy. First, the clustering procedure was applied to the initial pool of ion geometries, and an energy threshold of 300 meV was used to select the structures to be optimized with "tight" settings. Applying the "tight" instead of "light" settings does not change the energy hierarchy dramatically, but there are a number of structures that converge to a significantly different minimum on the PES, as shown in Figure 3-16 for the case of TC backbone conformation. Overall, the energy differences slightly decrease with the higher precision calculations, while the CCSs stay the same within 1 % (except for the outliers).
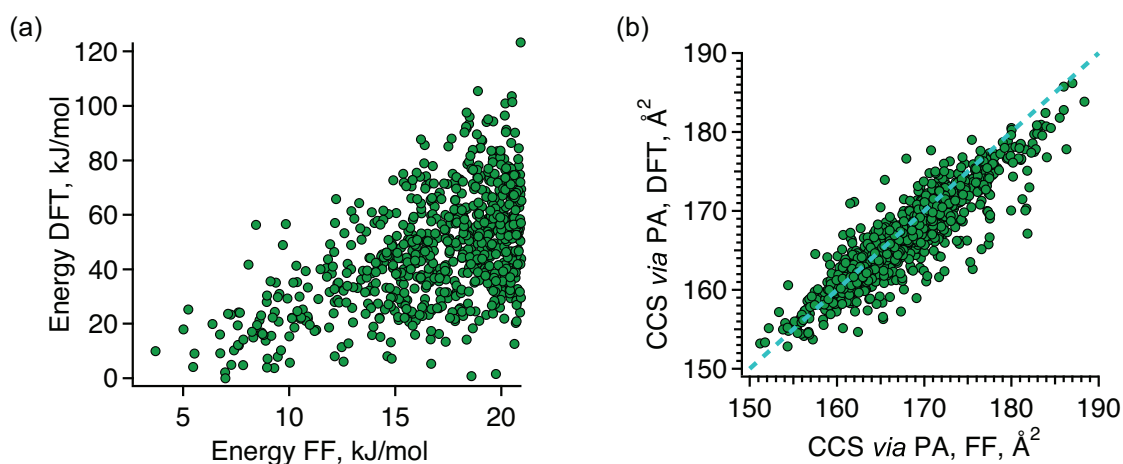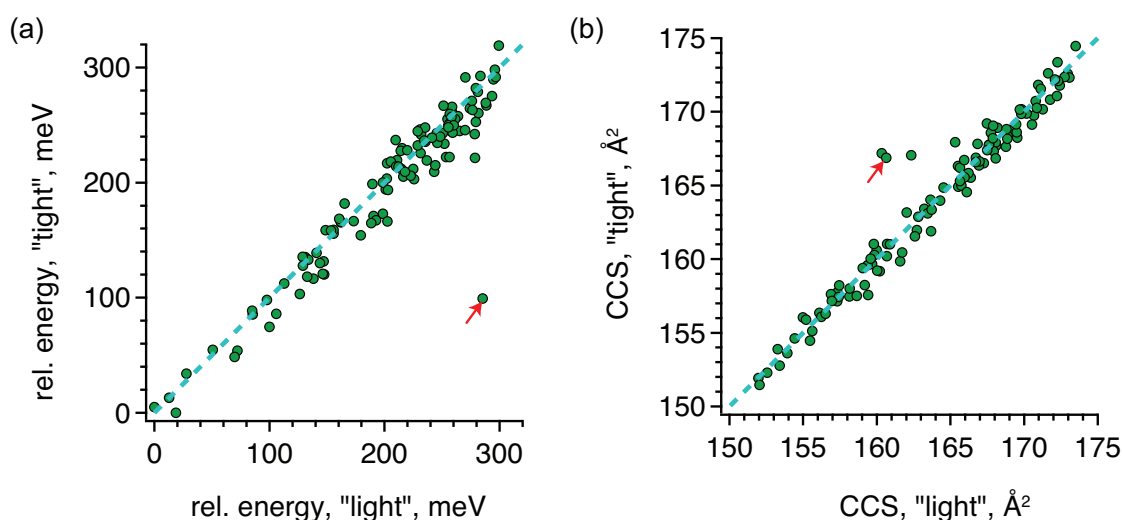

Figure 3-16. (a) The energies of all TC structures after the DFT optimization with the "tight" settings as a function of the energy obtained with the "light" settings. The dashed line corresponds to the equal values. An outlier is labeled with a red arrow. (b) The CCSs of all TC structures after the DFT optimization with the "tight" settings as a function of the CCSs obtained with the "light" settings, both computed with the TM. The dashed line corresponds to the equal values. An outlier, for which the energy changes drastically, is labeled with a red arrow.

From 429 structures optimized with the "tight" settings, we selected the lowest-energy conformations in each proline configuration (TT, TC, CT or CC, 57 in total) and computed their free energies and vibrational spectra in a harmonic approximation using the PBE functional and "tight" settings. Visual inspection of these spectra allowed us to choose the structures that best correspond to those we observed experimentally. We re-computed the spectra of all promising candidates using the PBE0 functional.

Changing the functional from PBE to PBE0 has an effect on the CCSs, the energy hierarchy and the vibrational spectra:

- The general trend is a slight compaction of the geometry upon re-optimization with a functional of higher accuracy (Figure 3-17(b)). We verified that this effect couldn't be explained by a different partial charge distribution: the CCSs with the charge artificially set to zero follow the same trend. We report here the CCSs calculated using PBE0 functional, as it is of higher precision.
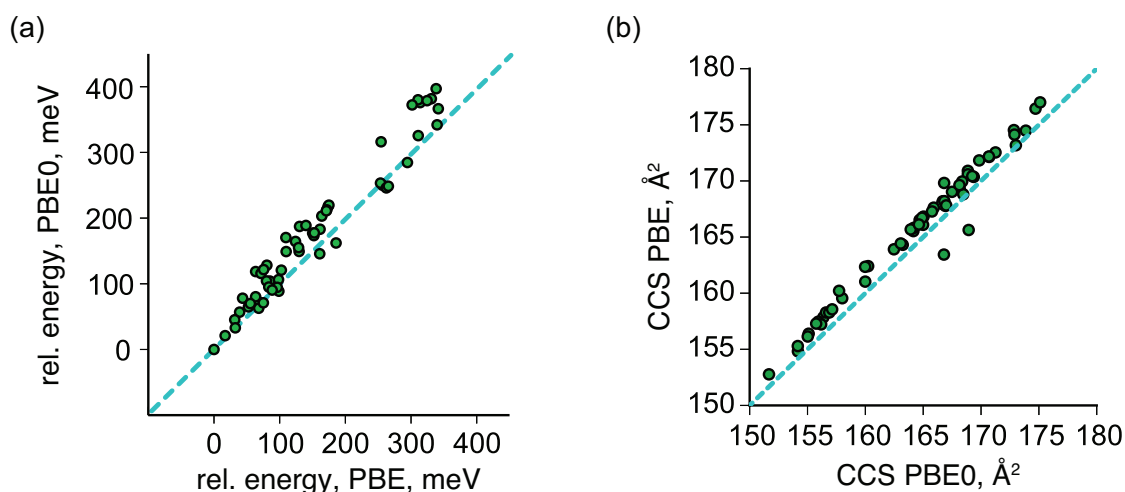
Figure 3-17. (a) The energies of 57 lowest-energy structures after the DFT optimization with the PBE0 functional as a function of the energy obtained with the PBE functional, both used with "tight" convergence settings. The dashed line corresponds to the equal values. (b) Change in the CCS values upon re-optimization of the 57 lowest-energy geometries with PBE0 functional instead of PBE. The dashed line of equal CCSs guides the eye.

- Energy hierarchies were compared for PBE and PBE0 functionals, both used with "tight" settings, for the 57 lowest-energy conformers of all structural types (Figure 3-17(a)). No outliers are observed, and the energy hierarchy is in general preserved, which confirms that relative energies computed with PBE functional predict well the energy differences calculated with PBE0 and, as a consequence, these two functionals share a set of low-energy conformers.

- Certain bands in the vibrational spectrum are more sensitive to the improvement in the level on theory than the rest of it. As a hybrid functional, PBE0 describes electron correlation more accurately than PBE, and thus the delocalized electrons of the phenylalanine side chain are described better. It

results in an improvement in the prediction of the position of the carboxylic OH-stretch relative to the arginine asymmetric stretch vibration (Figure 3-18).
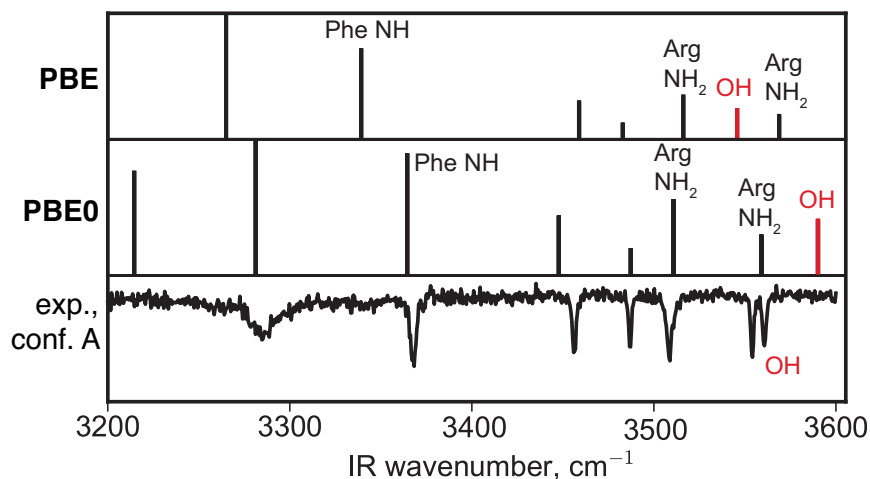
Figure 3-18. Vibrational spectra of a BK[1-5]$^{2+}$ low-energy conformer calculated with PBE and PBE0 functionals. The position of carboxylic OH stretch, which is tentatively assigned by the comparison between the electronically excited and ground states, changes the most.

The best-fitting spectra resulting from the simulations using PBE0 and the corresponding structures are shown in Figure 3-19. The calculated NH-stretch vibrational frequencies correspond remarkably well to the measured frequencies, including the hydrogen-bonded NH stretches. However, we observe a systematic shift between calculated and experimental frequencies of the free OH stretch vibration. Such a discrepancy can result from (i) the possible incomplete inclusion of electron exchange and correlation and (ii) the use of a single scale factor to globally account for vibrational anharmonicity and nuclear quantum effects.[339] The former issue is resolved by the use of the PBE0 hybrid functional with exact Hartree-Fock exchange. Provided a large data set of computed vs experimentally measured vibrational frequencies is available, the latter could be resolved by finding scaling factors for each vibrational mode by linear regression and subsequently using them for a system with unknown geometry. With this approach, it has been shown that these factors differ between NH and OH stretches.[340] In the absence of such a data set, we use a single scale factor (0.948 for PBE0 functional) to account for vibrational anharmonicity. In light of this, it is not surprising that the calculated OH vibrational band is shifted from experiment, since the scale factor is largely determined by the agreement for the more numerous NH stretch bands. Routes towards a correct simulation of the experimental OH stretch peak positions without the need for scaling are known – simulations that fully

include anharmonic effects and a quantum mechanical treatment of the nuclear motions. Examples are the multi-configurational time-dependent Hartree method[341] or the approximate thermostatted ring-polymer molecular dynamics method.[342] However, such methodology is not yet straightforwardly applicable to molecules of the size investigated here.

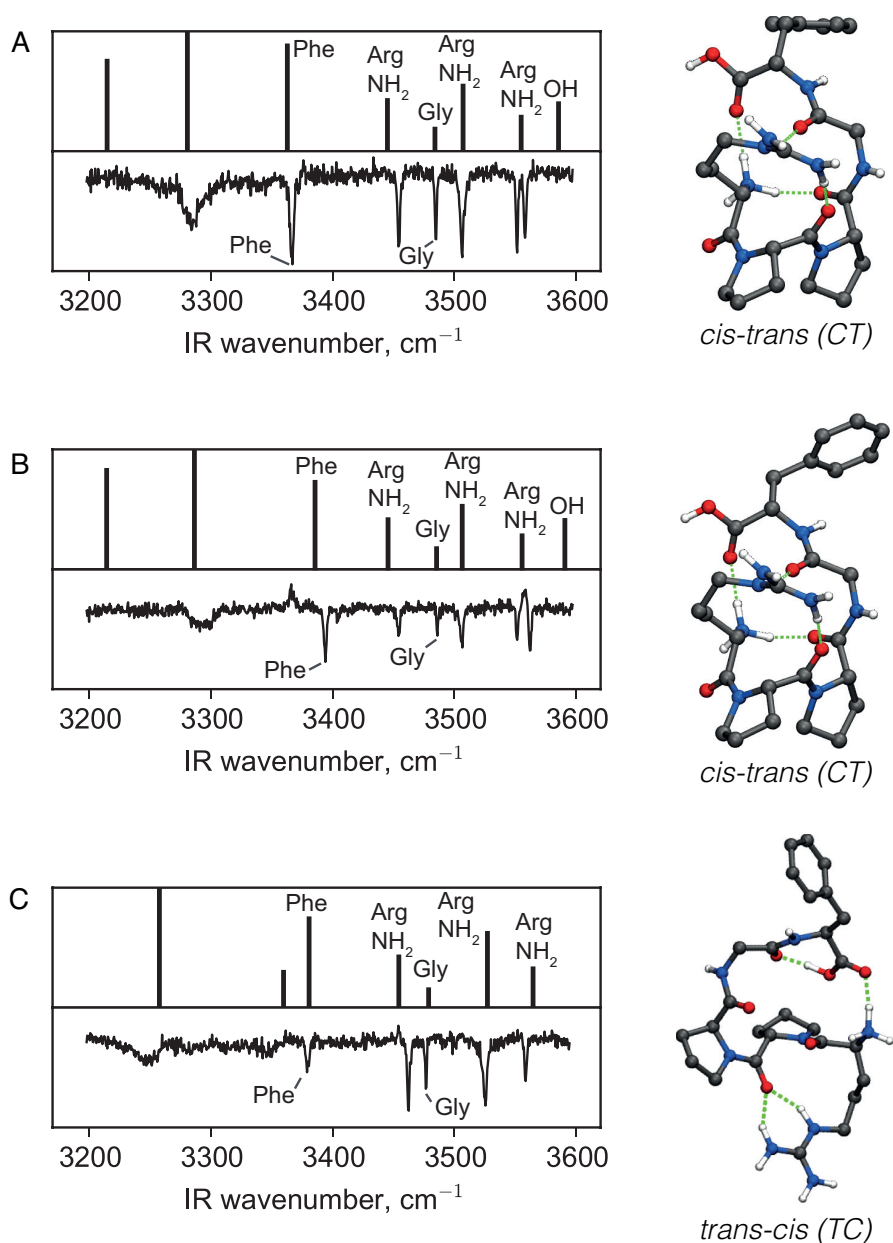Figure 3-19. Comparison between calculated and experimental spectra for conformers A, B and C, forming peak I. The corresponding structures are shown and the conformation of prolyl-peptide bonds is noted for each conformer. The NH stretch vibrations in the experimental spectra are identified via isotopic labeling. For the computed spectra all relevant local modes are labeled.

As shown in Figure 3-19, we were able to assign structures that fit to the spectral signatures that we found in peak I of the CCS distribution. Conformers A and B differ by the orientation of the phenylalanine ring, as the experimental data suggest. Conformers A and C have different configurations: in A the prolines are in a *cis-trans* configuration, while in C they are *trans-cis*. Despite the difference in structure, their calculated CCS values are very close: 169 Å² and 168 Å².

We measure ion mobilities at room temperature and then cool the ions in the cold ion trap. Depending on the rate of cooling, one might expect a certain degree of re-equilibration of the conformational distribution in this process. If the cooling is fast with respect to the isomerization rate between conformers, kinetic trapping can occur, preventing the preferred geometries at 300 K from converting to the low-energy, gas-phase structures at low temperature. In order to gain insight into the dependence of the energy ordering of the different conformations on the temperature of the molecule, we computed the energies of all 57 candidates with the PBE0 functional and then evaluated the free energy for a range of temperatures using a harmonic approximation. Figure 3-20 shows how the relative energies change within the CT, TC and CC categories for 3 cases: pure potential energy, free energy at 10 K, and free energy at 300 K. The importance of the free energy correction can be demonstrated using the example of the conformation that corresponds to the global minimum on the potential-energy surface (PES). It is a very compact CT structure with a CCS of 157 Å² and is not observed in the experiment. Indeed, at 300 K this conformer no longer represents the global minimum.

The energy difference at 300 K between conformer A and the global minimum is within 0.6 meV per atom, which is at the limit of what can be resolved at this level of theory. The global minimum at 300 K (GM 300 K in Table 3-2) is a CT structure that has the same hydrogen bonding pattern as A and B but which differs from them mainly by rotation of the phenyl ring (Figure 3-21). A number of other low-energy structures at 300 K with similar H-bonding patterns and CCS as A and B have been identified. We speculate that at 300 K the structures of the CT type are close to the global minimum and are rapidly interconverting, while after cooling the ions in the trap we freeze out mainly conformer A and traces of conformer B. The higher the number of similar structures available around the global minimum, the higher is the volume of conformational space corresponding to this basin at 300 K and, accordingly, the higher is the probability to find the molecule there.
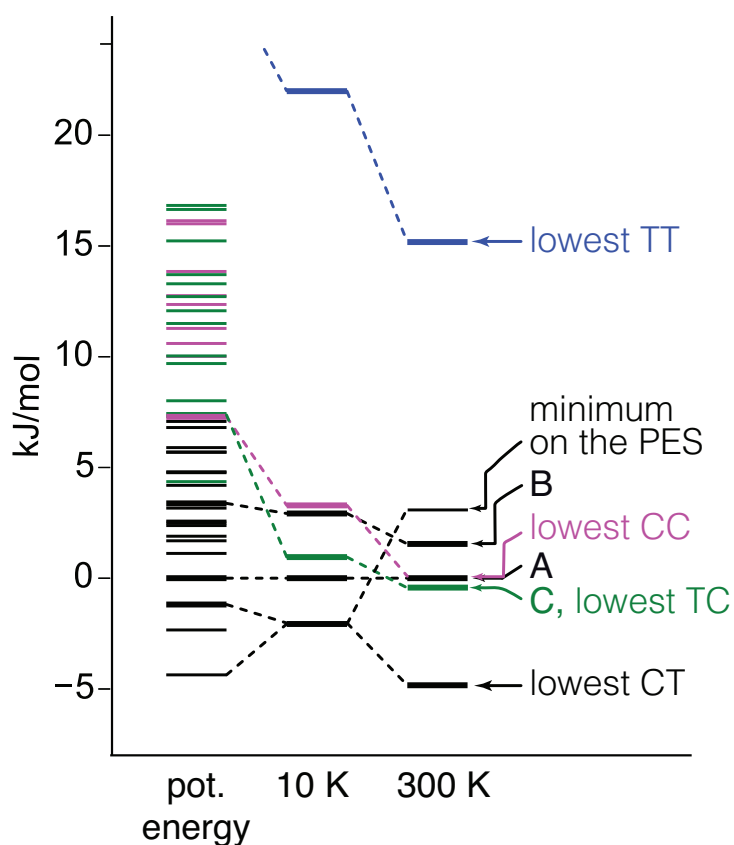
Figure 3-20. On the left: relative potential energies within CT (black), TC (green) and CC (magenta) configurations. On the right: the change in relative energies of the main conformers for 3 cases: pure potential energy, free energy at 10 K and free energy at 300 K. A, B and C label the experimentally observed conformations. The lowest energy TT structure (blue) is given for a reference. The lowest-energy geometries at 300 K for TC, CT and CC configurations are labeled as well.
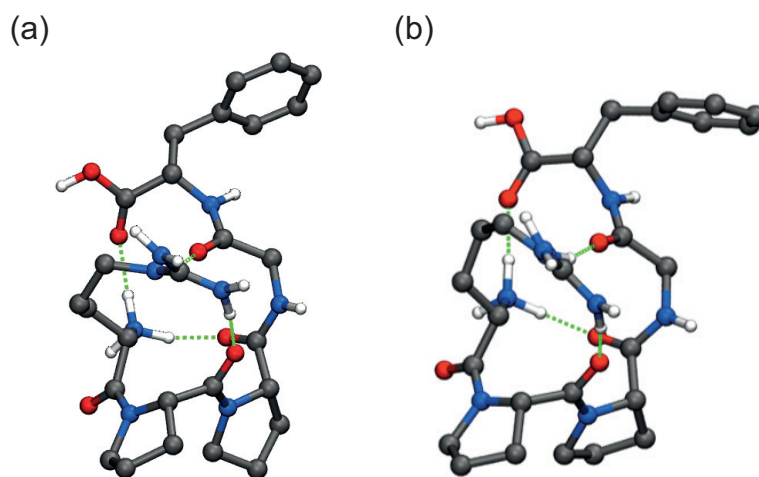
Figure 3-21. (a) The structure of experimentally observed conformer B, (b) the structure of the free energy global minimum at 300K. Both conformers fall into CT type and their hydrogen bond pattern is identical.

Conformer C is the lowest-energy structure in the TC category at 300 K. In general, we observe in the experiment the low-energy geometries for CT and TC types but not of the CC type. Further computations of energy barriers between different conformers are needed to fully understand why only a fraction of the available low-energy structures are observed experimentally. It could also lead to a better understanding of the interplay between the collisional cooling rate and the isomerization rate.

3.3.2 Extended kinetically trapped structures

While there is a relatively well-established (though computationally expensive) procedure for searching for the global minimum on a potential energy surface, identifying kinetically trapped conformations is challenging. One cannot rely on the energy criterion, because the kinetically trapped structures lie, by definition, high in energy and are separated from other conformational basins with even higher energy barriers. In this case, experimental constraints play a key role in guiding quantum-chemical calculations. One of them is the CCS. In the pure IM studies CCSs alone are often used to assign the structural motifs to certain features in the IM distribution. To test this approach, we computed CCSs of all geometries in the initial pool of structures obtained *via* DFT optimization with "light" settings and compared the distribution of structures over different CCSs and potential energies as a function of the backbone type (TC, CT, TT or CC, Figure 3-22).

These calculations lead to several important observations:

- The statement that the CC conformers are more compact than the other backbone types does not hold for this system. On the contrary, based only on the plots presented in Figure 3-22 one would assign the extended structures to the CC type.

- The TT structures are by far the highest in energy in the gas phase, which is opposite from the solution phase behavior.

- The range of CCSs achievable for the system is broader than the one measured experimentally, and much more compact structures are predicted to exist based on the PES shape. We show below, however, that taking into account the temperature at which the IMS experiment is performed (300 K) changes the energy distribution and favors larger CCSs.
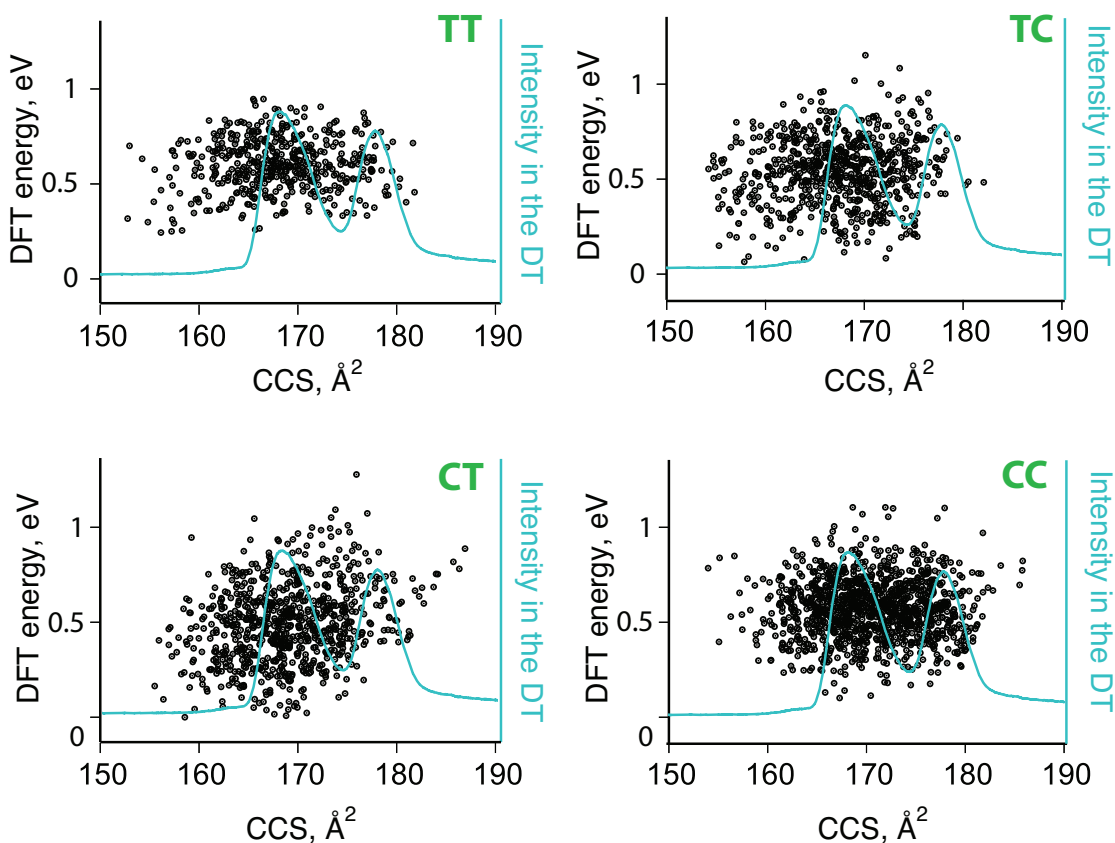
Figure 3-22. Potential energy after DFT optimization as a function of CCS for the structures of different backbone conformation. All energies are relative to the lowest-energy structure, which belongs to the CT type.

Most importantly, the distributions in Figure 3-22 demonstrate that assigning the structures or even structural families based on the CCS alone is hardly possible. The PES of this unstructured BK fragment is relatively flat and no structural motif is favored over the others in a definite way. We thus continued the analysis within every backbone type in order to use the spectroscopic signatures to verify the theoretical prediction.

We considered only those structures that fall under peak II in the arrival-time distribution, as shown in Figure 3-23(a). The energy threshold for a structure to be considered was kept as high as possible: 48 kJ/mol (0.5 eV) from the lowest-energy conformer in each type. We calculated vibrational spectra at the PBE+MBD level for all structures that satisfied these two conditions (135 in total).
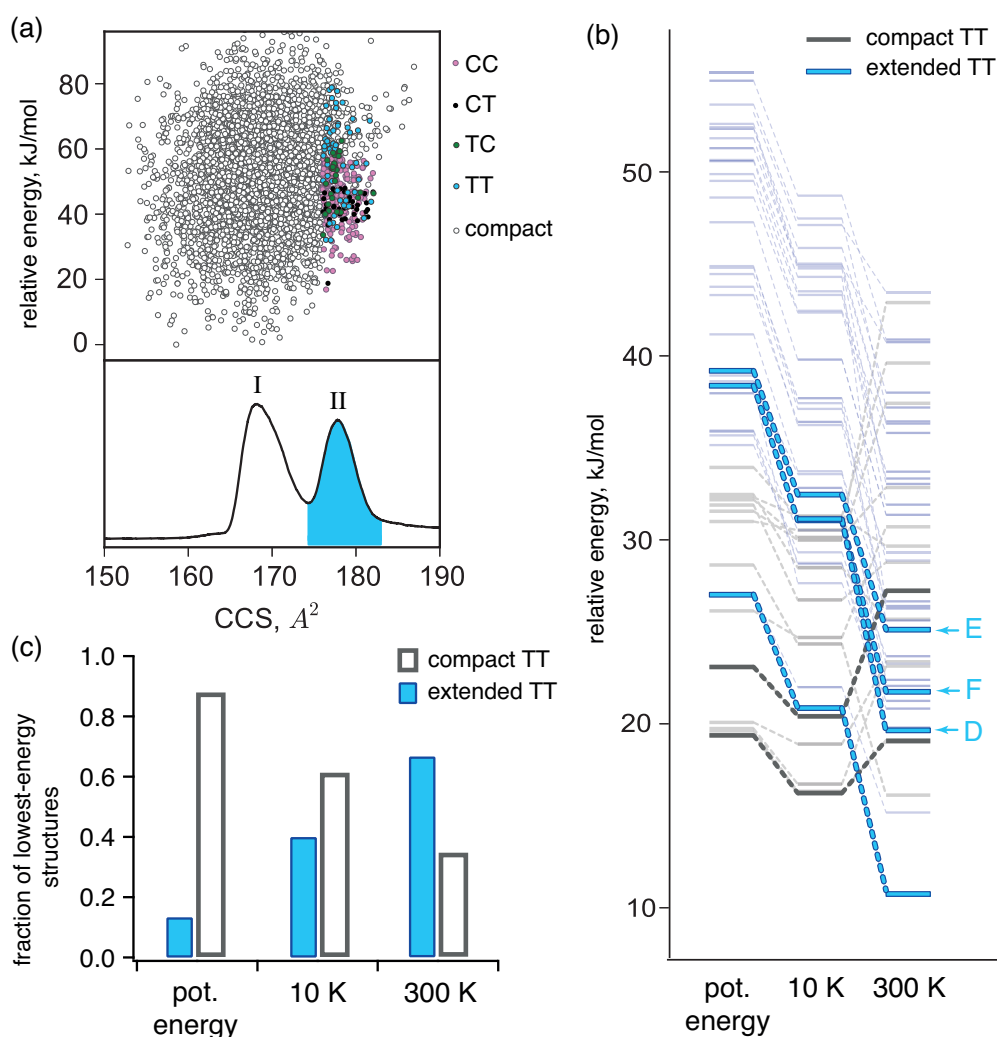
Figure 3-23. (a) Distribution of relative potential energies of considered conformers as a function of CCS. Those conformations for which the vibrational spectra were computed are shown in color. The rest of the structures are shown in grey. (b) Energy hierarchy for the lowest-energy structures of TT type. The extended structures with CCS higher than 176 Å^2 are shown in blue, the compact ones – in grey. The structures that are found in the experiment and several representative low-energy structures are highlighted to show the general trend. (c) The ratio between compact and extended (CCS higher than 176 Å^2) structures among 15 lowest-energy conformers without and with temperature correction.

This enabled a second selection step, apart from the CCS cutoff, based on spectroscopic information:

1. In the IR spectrum of extended structures (see Figure 3-7) there are two bands above 3570 cm^{-1}. The only vibration in this molecule that can have such high frequency is the OH stretch, suggesting that there at least two conformers under peak II of the drift-time distribution and each of them has a free OH stretch.

2. In the IR spectrum of conformational family II there is a unique, intense band at 3538 cm^{-1}, which is too high for a backbone NH vibration and thus corresponds to

the side chain of arginine.[343] To decide if this side chain interacts with other parts of the molecule, we compare its spectroscopic signature with those of conformers A, B and C. Their spectra show that the asymmetric stretch of a *free* η-NH$_2$ of arginine appears at a frequency higher than 3540 cm^{-1} while the symmetric stretch occurs between 3450 cm^{-1} and 3500 cm^{-1}. Since the intense band in the spectrum of peak II (see Figure 3-7) is slightly lower than this, we conclude that the extended conformers have no free η-NH$_2$ groups, but a least one weakly interacting with other parts of the molecule.

3. The free ε-NH stretch of arginine should appear at 3465 cm^{-1}.[344,345] The absence of this band in the spectrum suggests that the ε-NH of arginine in all the extended conformers is hydrogen bonded.

4. The strong bands at 3362 cm^{-1} and 3303 cm^{-1} are most probably the spectroscopic signature of the free NH$_3^+$ group or one weakly bound to a phenyl ring.[346]

5. Based on the comparison with the spectra of conformers A, B and C, the phenylalanine NH stretch might interact with the phenyl ring, while the NH stretch of glycine appears to be free.

With these considerations we aim to rationalize the comparison between experimental and simulated spectra that otherwise is easily prone to incorrect or non-precise assignment. It is necessary to analyze the spectroscopic features manually and to perform isotopic labeling. After applying the criteria stated above to the simulated spectra and molecular geometries, only structures of type TT remain, meaning that both prolyl-peptide bonds are in the *trans* configuration. We re-computed the IR spectra of those structures that satisfy the above criteria with the PBE0 functional, which further improved the agreement between simulated and experimental vibrational spectra (Figure 3-24(a)).

Figure 3-23 (b) and (c) explain why we observe elongated TT structures in the ion mobility experiment. They are more stable at 300 K than the compact structures, while the latter become more energetically favorable at 10 K. At 300 K the elongated structures traverse the drift tube, rapidly interconverting between the close minima on the free energy surface. Then they are transferred to the cold ion trap, where the final structure at 10 K results from the interplay between collisional cooling and isomerization. The fact that we observe spectroscopically the elongated structures suggests that some of their features are preserved in the cooling process. Kinetic trapping could occur twice during their

transition from solution to our cold ion trap: in the ion source and in the cryogenic ion trap. We observe evidence of some kinetic trapping in both cases.
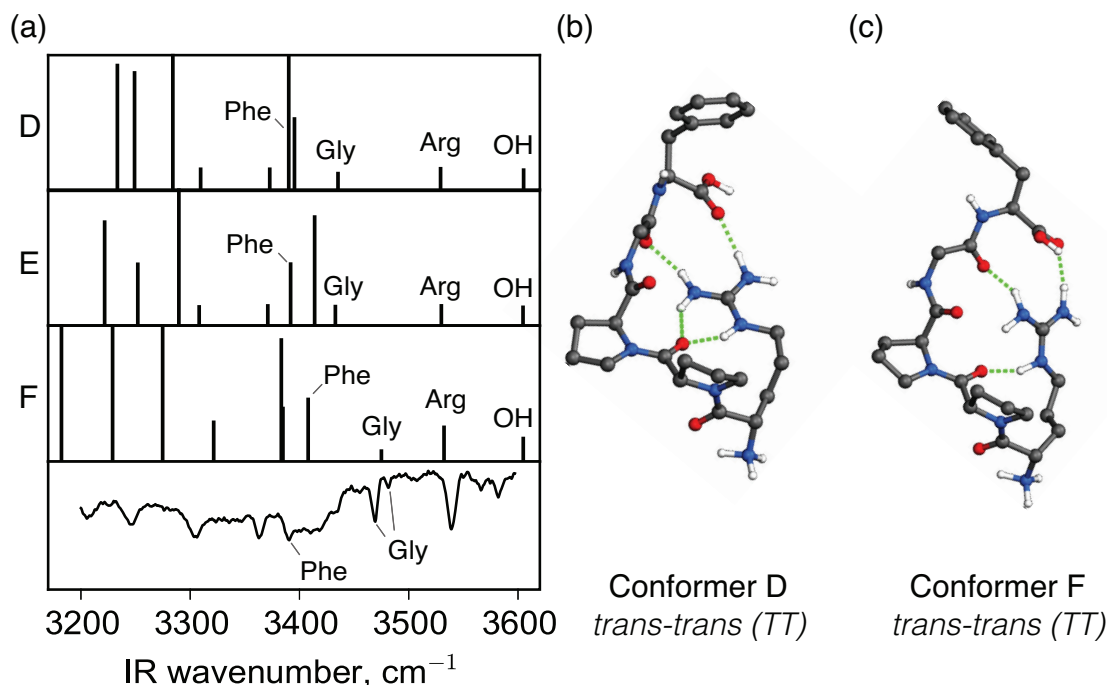
Figure 3-24. (a) Calculated vibrational spectra of conformers D, E and F compared to the experimental spectrum of peak II of the arrival-time distribution; (b), (c) the corresponding structures. Structure of conformer E is visually almost indistinguishable from the conformer D and is not shown.

3.4   Bradykinin 1-5 in solution

Numerous NMR studies of the full bradykinin molecule indicate that in aqueous solution it mainly adopts an all-*trans* conformation,[180,184] which is consistent with the *trans* bond being slightly more stable than *cis* in polar environments due to its higher dipole moment.[180] Two receptors of bradykinin are known,[178] B1 and B2, and in the complex with the B2 receptor bradykinin is in the all-*trans* form.[184,185] If the medium becomes non-polar, alternative backbone conformations involving *cis* prolyl peptide bonds are more competitive.[188]

We observe the same type of behavior for BK[1-5]$^{2+}$, which acts *in vivo* as an inhibitor of thrombin. An X-ray structure of the complex demonstrates that all peptide bonds in BK[1-5] remain *trans*.[200] A part of the thrombin complex with the first four amino acids of BK[1-5] sequence reconstructed from the X-ray measurements is shown in Figure 3-25.
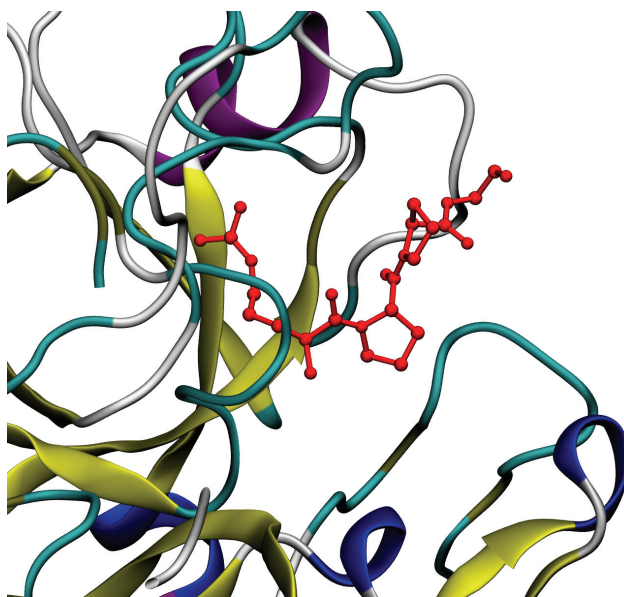
Figure 3-25. A part of thrombin – BK[1-5] complex reconstructed from a pdb structure file.[200] The first four amino acids of BK[1-5] are visible and shown in red, both prolines are in *trans* conformation.

We performed NMR measurements in a CD_3OH : H_2O mixture, which reproduces the solvent used for the ESI experiments (Figure 3-26). Peaks were assigned using COSY and TOCSY 2D ^{1}H spectra with water suppression, and the assignments are supported by comparison with spectra acquired in DMSO, where no background water signal hinders the cross-peaks. No residual secondary structure is identified based on the Hα secondary chemical shifts, which are close to zero for all residues but N-terminal arginine (Table 3-1).[347,348]

Table 3-1. Chemical shifts of bradykinin 1-5 in $CD_3OH : H_2O$ 50:50 solution compared to the random coil values according to Wishar et all.[348] The Hα protons that reflect the secondary structure are highlighted.

Residue number	Residue and atom	Measured ^1H chem. shift, ppm	Random coil ^1H chem. shift, ppm	Secondary chem. shift, ppm
1	**Arg Hα**	**4.34**	**4.65**	**-0.31**
1	Arg Hβ	1.93	1.81, 1.81	+1.12
1	Arg γCH$_2$	1.72	1.67, 1.67	+0.05
1	Arg δCH$_2$	3.2	3.21, 3.21	-0.01
2	**Pro2 Hα**	**4.79**	**4.73**	**+0.06**
2	Pro2 Hβ	2.42	2.31, 1.91	+0.11
2	Pro2 γCH$_2$	2.01	2.01, 2.01	0
2	Pro2 δCH$_2$	3.78, 3.52	3.6, 3.6	+0.18, -0.08
3	**Pro3 Hα**	**4.42**	**4.44**	**-0.02**
3	Pro3 Hβ	2.26	2.29, 1.94	-0.03
3	Pro3 γCH$_2$	2.06	2.02, 2.02	+0.04
3	Pro3 δCH$_2$	3.68, 3.83	3.63, 3.63	+0.05, +0.2
4	**Gly Hα**	**3.93**	**3.96**	**+0.03**
4	Gly NH	8.34	8.33	+0.01
5	Phe NH	7.99	8.3	+0.31
5	**Phe Hα**	**4.62**	**4.62**	**0**
5	Phe Hβ	3.2, 3.03	3.14, 3.04	+0.06, -0.01
5	Phe ring	7.24, 7.33, 7.29	7.28, 7.38, 7.32	-0.04, -0.05, -0.03

ROESY 2D ^1H NMR spectrum shows that 82% of molecules in a 50:50 water : methanol mixture adopt the *trans-trans* configuration (Figure 3-26). Three alternative conformers do not have a fully resolved spin system and thus it is difficult to identify them.
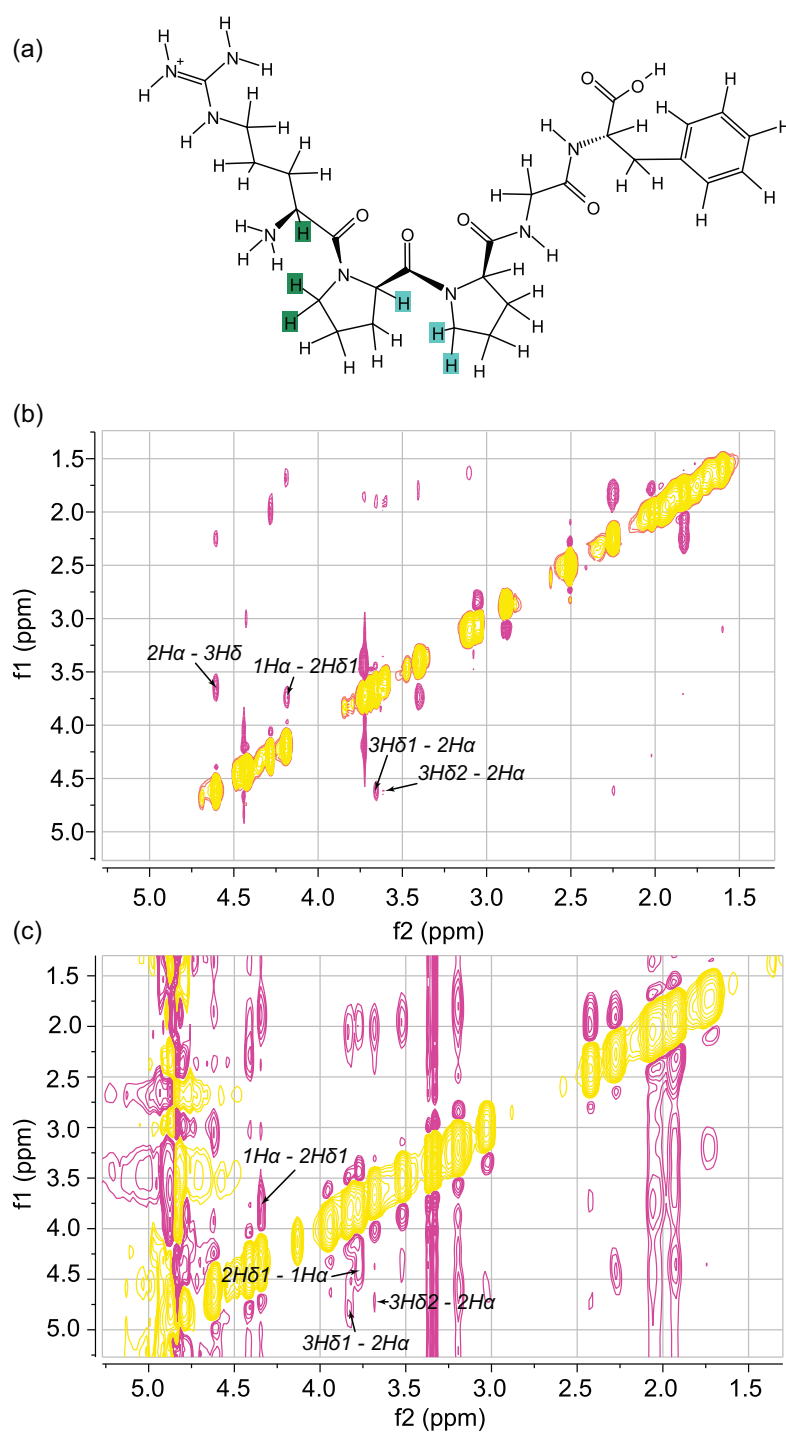
Figure 3-26. Determination of the backbone conformation of BK[1-5] in solution by NMR. (a) The structure of BK[1-5] with explicit hydrogens. The hydrogens, which are spatially close in *trans* conformation and distant in *cis* are highlighted for each prolyl peptide bond. ROESY spectra of BK[1-5] in (b) DMSO solution, (c) CD₃OH : H₂O 50:50 solution with mixing time of 400 ms. The background is attributed to incomplete water and CD₃OH signals suppression. The peaks that correspond to the *trans* conformation are shown.

3.5   Overview of the conformational space of bradykinin 1-5

As shown in section 3.4, in the methanol : water 50:50 mixture both prolyl-peptide bonds of BK[1-5] are mostly in the *trans* conformation. Upon electro-spraying from the same water : methanol mixture, we observe a kinetically trapped conformational family in which all the bonds remain *trans*. These TT conformers are folded in such a way that the N-terminus of the molecule remains free, while the arginine side chain interacts with the C=O groups of the backbone. The broad ion-mobility peak might reflect the disordered behavior in solution. Inside the drift tube at room temperature the close conformations might be rapidly interconverting, while at 10 K in the conditions of the cold ion trap the structure becomes well defined. We do not observe the lowest-energy structures of the TT type in the experiment, but rather the ones that lie quite high in energy. This means that the mechanism of kinetic trapping involves not only *cis-trans* isomerization, but also H-bonding re-arrangement, which can have high-energy barriers as well (Figure 3-27).
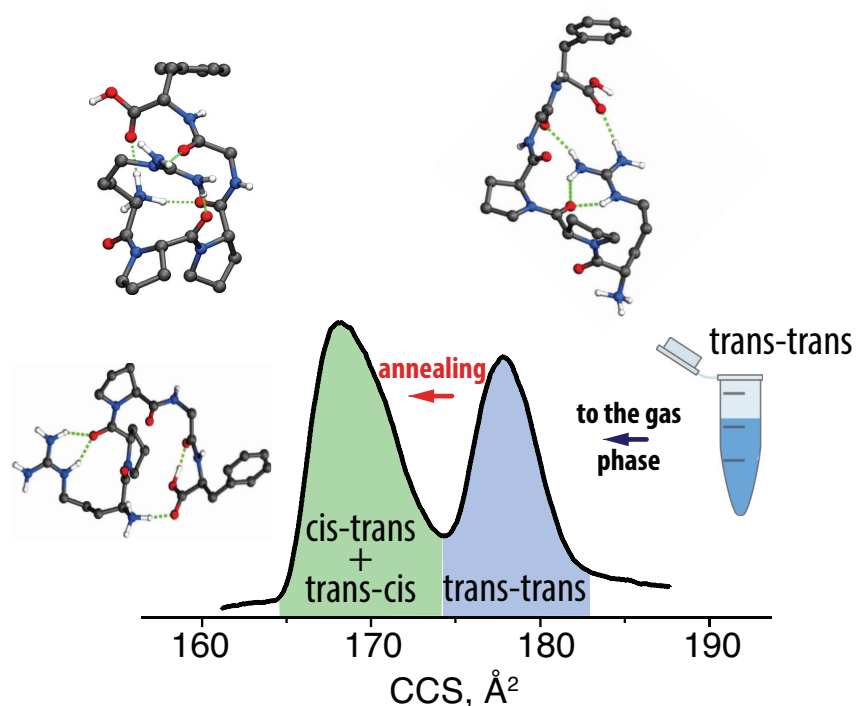
Figure 3-27. Summary of the findings. Most of the molecules are trans-trans in solution; this structural feature is preserved in the kinetically trapped species, but can be disrupted *via* annealing in the gas phase, which leads to a change in the conformation of one of the prolines.

Upon collisional activation, TT conformers of BK[1-5]$^{2+}$ (i.e., the extended structures in the IM distribution) convert to TC and CT structures (the compact ones). The relative abundances of these conformers at different degrees of activation observed by cryogenic ion spectroscopy (see Figure 3-8) suggest that first the TT to TC transition occurs, followed by TT to CT. The alternative pathway via CC conformers is unlikely as we do not find any trace of CC conformers, even though their free energies are close to those of conformers A and C. Further investigation of conversion paths between all types of structures will shed light on this issue. One explanation might be that the barrier height between CC and the rest of the conformational space is significantly higher than between TT and CT or TT and TC, so that the molecule fragments before sampling the CC states.

In all cases except TT, the N-terminus and arginine side chain can both form H-bonds, which stabilize the molecule compared to TT. The external hydrogen bonds have to be substituted by internal ones upon the removal of water. This is why in the gas phase even the lowest-energy TT structure is significantly less stable than TC, CT and CC (Table 3-2). These latter three types have compact low-energy structures with CCSs falling within peak I in the drift-time distribution. Low-energy conformers of the TT type tend to be more extended, with CCSs close to peak II in the drift-time distribution. Overall, the calculated CCSs presented in Table 3-2 agree well with the values that we measure (Figure 3-28), given that the experimental error is estimated to be around 2 Å^2. The compaction in the gas phase corresponds to the effect observed for BK adhered to the surface of a micelle.[349]

Figure 3-28. Comparison between the computed and experimental CCS for the three conformational families identified in the present study. The green line corresponds to conformations A and B, the orange one – to C, the blue one – to D. The experimental error is shown. See Table 3-2 for more details.

Table 3-2. Comparison of the major conformers identified in the present study. GM denotes the global minimum on the corresponding energy surface.
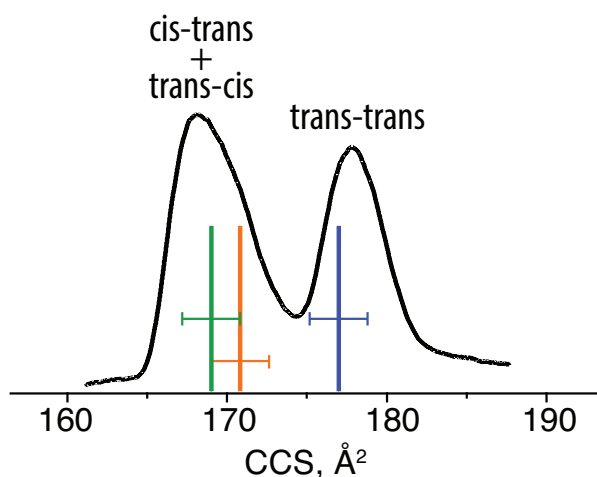
Conf.	Prolyl-peptide bond conformation	Potential energy, meV[a]	Free energy at 10 K, meV[a]	Free energy at 300 K, meV[a]	CCS experiment, Å²	CCS computed, Å²
GM (pot)	CT	-45	-21	32	-	157
GM (10 K)	CT	12	-24	-36	-	165
GM (300 K)	CT	-12	-21	-50	-	168
A	CT	0	0	0	167, 168	169
B	CT	35	30	16	167, 168	169
C, lowest TC (10 K and 300 K)	TC	76	10	-5	167, 168	171
Lowest TT (300 K)	TT	280	216	111	-	175
D	TT	394	322	204	176	177
E	TT	406	336	260	176	176
F	TT	398	322	225	176	177
Lowest CC	CC	75	34	1	-	169

[a] Relative to conformer A

3.6 Conclusions

The combination of ion mobility and cryogenic ion spectroscopy has allowed us to characterize all major conformations of BK[1-5]$^{2+}$ in the gas phase. Our calculations reveal that three very different conformational families (*cis-trans*, *trans-cis* and *cis-cis*) have close free energies and CCSs in the gas phase, and, by combining the experimental spectroscopic information with first-principles simulation data, we were able to assign them. We determine that the lowest-energy gas-phase structures of the peptide feature backbone conformations with the two prolyl-peptide bonds in either *cis-trans* or *trans-cis* state. Both of them fall within the more compact ion mobility peak at 167 to 170 Å².

Kinetic trapping can occur at different points in the "life cycle" of the ion in our instrument. First, when the ions are initially desolvated in the electrospray, the conformations coming out of solution can be kinetically trapped behind barriers that result from internal hydrogen bonding and separation in *cis-trans* states of the prolyl-peptide bonds. Second, if cold-ion spectroscopy is used, kinetic trapping can occur when cooling the molecules. In this case, when the collisional cooling rate is faster than the isomerization rate, the population of conformers that are higher in energy than the global minimum can be trapped behind isomerization barriers.

We have identified the kinetically trapped conformers of BK[1-5]$^{2+}$ as all-*trans*, which is the predominant backbone conformation type in solution. This provides the basis for further gas-phase studies to investigate the isomerization barriers between the prolyl-peptide bond types CC, CT, TC, and TT and to study the extent of micro-solvation that is required to render TT more stable than the alternative backbone types. The width of the observed IM peak and the spectroscopic signature of the kinetically trapped conformations suggest a wide variety of structures, which is in agreement with the intrinsic disorder in solution. In stark contrast to this, we have shown that the stable, gas-phase *cis-trans* and *trans-cis* conformers appear to represent well-defined minima. This is consistent with the notion that disordered peptides collapse to particular secondary structures in non-polar media such as cell membranes.

3.7 Additional experimental details

Bradykinin[1-5] (RPPGF, Trifluoroacetate salt, Bachem) was purchased and used without further purification. Peptide solutions were prepared in a 49:49:2 mixture of water : methanol : acetic acid with a peptide concentration of 50 μM. In the 2+ protonation state, the extra protons are believed to reside on the arginine side chain and on the N-terminal amine.

Chapter 4.
The case of triply protonated bradykinin

In this chapter we study triply protonated bradykinin in the gas phase and qualitatively compare the results to its solution-phase behavior. Field-asymmetric ion mobility spectrometry is used to separate the conformational families and identify kinetically trapped structures, which were characterized using double-resonance cold-ion spectroscopy. We assign the bands in the vibrational spectra to particular oscillators in the molecule using nitrogen isotopic labeling. We also show how carbon isotopic labeling of the phenylalanine side chain can be used to obtain additional information about the molecular geometry and record vibrational spectra of the lowest-energy conformers of BK^{3+}, which appears challenging otherwise. Part of the results presented here have been published previously.[350]

4.1 Introduction

In Chapter 3 we focused on the bradykinin 1-5 fragment and established that in the most abundant +2 charge state the conformational distribution in the gas phase contains kinetically trapped structures that have the same backbone conformation as the solvated molecule, while the low-energy structures are also populated. Here we study the full bradykinin (BK) sequence (Figure 4-1). It consists of nine amino acids, three of which are prolines. By analogy with BK[1-5] and in agreement with previous studies,[173] prolines should contribute to the conformational heterogeneity of the peptide.
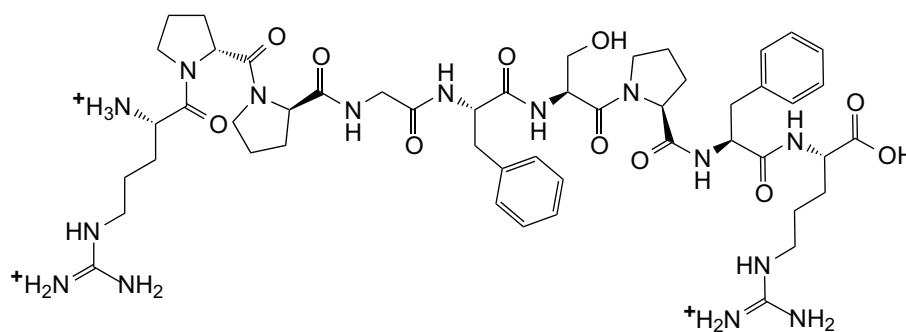
Figure 4-1. Chemical structure of triply protonated bradykinin.

In solution and *in vivo*, BK has been studied extensively, and its conformational preferences are mostly understood. The life cycle of BK in a living organism begins with it being cut out of the sequence of the precursor peptide kininogen by specific proteases in the blood.[351] BK binds to the G-protein coupled receptors B1 and B2, which is linked to inflammation and blood pressure regulation.[11,178,352] The initial BK degradation in human plasma results in the loss of the last amino acid producing des-Arg^9-BK.[353] Both this product and BK are further proteolitically cleaved by an angiotensin converting enzyme and result in the fragment BK[1-5], our study of which was described in Chapter 3.[354] In addition to these changes in its sequence, BK experiences drastically different environments: from the dominantly aqueous solution of the bloodstream and interactions with a variety of membranes of different composition and charge (e.g. different surface net charge depending on the cholesterol content) to the binding sites in the respective receptors or proteases. The structural properties of BK in all of these situations are crucial for understanding its function.

In all media employed for the NMR studies, the proline conformation was determined to be mostly *trans-trans-trans*, though the exact percentage of the minor conformations involving *cis* configuration of the peptide bond differs from case to case (Table 4-1).

In the gas phase, BK can form three charge states. Singly protonated BK was investigated using MALDI and IMS as early as in 1996.[190] The peptide was found to adopt a globular conformation, which appears as a single peak in the IM distribution, independent of the temperature. Doubly protonated BK, as the most abundant charge state produced by ESI, has been studied extensively. Its IM distribution consists of two well-resolved peaks,[275,355] while FAIMS combined with H/D exchange was able to separate four different conformations.[191]

Table 4-1. Dependence of percentage of the *cis* conformation in any proline in BK sequence on the experimental conditions, such as the solvent and interaction with other substances. An unusually high ratio of *cis-trans* isomerism is observed if serine in position 6 is substituted with glycine.

Experimental conditions	Percent of *cis* conformers
water	10 %[158,356]
pH titration in water	10 %[356]
DMSO	<15 %[180]
lyso phospatidylcholine micelles	<10 %[180]
90:10 dioxane : water	< 8 %[180]
99:1 dioxane : water with calcium	minor[181]
water in presence of polyphenols	10%[182]
water in presence of lyposomes	minor[183]
[6]Gly-BK	35 %[158]

Upon improvement of IM resolution and sensitivity, it became possible to resolve three main (labeled A, B and C)[355] and up to 10 minor conformations of triply protonated BK.[139,196] For this charge state, Clemmer and co-workers showed that the conformational distribution produced in the ESI process is different from the quasi-equilibrium in the gas phase, and thus kinetic trapping takes place.[139] They then estimated the energy barriers between the conformational families[74] and suggested that the initial IM distribution depends on the solvent used for the ESI.[357] Moreover, three major conformational families were assigned to certain conformations of the backbone in relation to *cis-trans* isomerism (Table 4-2).[173] Later Wysocki and co-workers tentatively assigned a minor population within peak C to *cis-cis-trans* based on the low-energy CID spectra, DFT simulations and N-methyl alanine substitutions.[70] Both Clemmer's and Wysocki's groups agree that the conformational families within peak C are mainly produced in the gas phase, while A and B are kinetically trapped.

Table 4-2. Isomer forms of proline in BK[3+] structures suggested in the literature.[70,173]

BK conformer	Pro[2]	Pro[3]	Pro[7]
A	*cis*	*cis*	*cis*
B	*cis*	*trans*	*trans*
C major	*trans*	*trans*	*cis*
C minor	*cis*	*cis*	*trans*

In order to relate the gas phase conformations of BK to its solution phase structures Russell and co-workers recorded an IM-MS map of hydrated ions in a cryogenic drift tube.[82]

They observe only the +2 charge state of BK at a capillary temperature of 340 K, while higher capillary temperatures and full desolvation are required to produce appreciable amounts of the +3 charge state, and only low abundances (<1 %) of water clusters of BK^{3+} are present. The authors conclude that formation of triply protonated BK is "critically tied to the desolvation thermodynamics".[82]

Comparison between the data presented in the Table 4-1 and Table 4-2 suggests that the kinetically trapped conformations of the +3 charge state do not directly correspond to the major backbone conformation in solution. Do they instead represent minor conformations that cannot be fully resolved with NMR? Why does the *cis-cis-cis* conformer, which seems to be the least populated in solution, form the most abundant kinetically trapped structure (conformational family A)? The fact that it is kinetically trapped means that it is related to the ESI process and ultimately to solution, but in which way and where is it formed? Does the +2 charge state reflect the conformational distribution measured in the NMR studies? Alternatively, can there be another structural element, which is preserved in the ESI due to a high barrier on the PES, aside from proline *cis-trans* isomerization? We aim to address these questions by careful examination of the conformational space of BK in the gas phase.

In this work we use FAIMS as a separation technique, which has been shown to be partially orthogonal to the IMS.[358,359] This research is inspired by a study performed in our laboratory that demonstrated the existence of a kinetically trapped conformational family of the doubly protonated BK using FAIMS and revealed its UV spectra.[195] We implement a similar technique to study the +3 charge state of BK and perform direct comparison with the IM data. We then continue by obtaining IR spectroscopic fingerprints of the most abundant conformations and characterize them experimentally using isotopic labeling and vibrational spectra in the electronically excited state.

4.2   Ion mobility separation of BK^{3+}

Ion mobility studies of triply protonated bradykinin (BK^{3+}) by Clemmer and coworkers reveal three primary conformational families in the gas-phase, which they call A, B and C, and several elongated structures that each represent less than 5% of the population.[139] Under normal conditions, the population distribution among the three main families was found to be 22:31:31. Under conditions in which they anneal the conformer populations by collisional activation to a "quasi-equilibrium" (QE) distribution, this ratio

changes to 1.8:16:80. They obtain the same ratio if they first isolate a single conformation and then subject it to the annealing process.

We reproduced these experiments in our laboratory. The IMS distribution of BK^{3+} electrosprayed from a 50:50 dioxane : water mixture is shown in Figure 4-2(a). It consists of three major peaks and one minor at 279 Å^2, and we do not observe elongated conformers at CCSs larger than 310 Å^2.
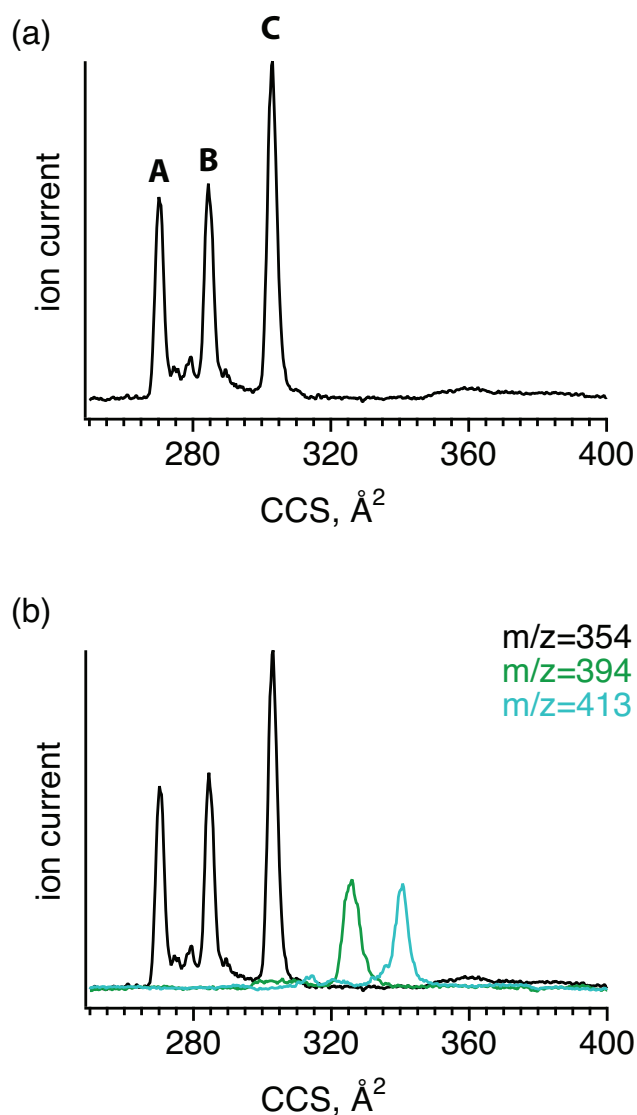

Figure 4-2. (a) IM distribution of BK^{3+} recorded upon ESI from a dioxane : water 50:50 mixture. (b) The same distribution compared to the IM distributions recorded on higher mass-to-charge ratios.

Apart from the expected ion current of the BK +2 and +3 charge states, we find ion signal on the mass-to-charge ratios higher than that of BK^{3+} and record the corresponding

IM distributions (Figure 4-2(b), the calibration is performed assuming +3 charge state for these species). Upon selection and subsequent collisional activation, the ions at m/z=394 and m/z=413 first increase in intensity due to improved transmission and then, in harsher conditions, interconvert completely to the stable gas-phase conformational distribution of BK^{3+}, as shown in Figure 4-3. We conclude that solvent adducts or multimers of BK are observed, and their charge state is not lower than +3. The mass-to-charge ratio 413 corresponds to a complex of BK^{3+} with two dioxane molecules. Given the chemical structure of dioxane, which can be also described as [6]-crown-2, it can potentially form a stable non-covalent complex with the protonation site on the peptide. However, the mass-to-charge ratio 394 cannot be readily attributed to any solvent adduct. We attempted to observe these adducts in a commercial LTQ Orbitrap Elite ETD (Thermo Fischer), but we find neither the m/z=394 nor m/z=413, probably, due to harsh conditions in the ion source. Further detailed study of this issue is necessary, as it might shed light on the last stages of the ES process.

In general, we find that the IM distribution of BK^{3+} is very sensitive to the source conditions. Typically, tuning the source voltages for the "gentle" conditions decreases the total ion current as the ion transmission deteriorates. We thus refrain from reporting the relative intensities of the peaks in the source distribution and only rely on their relative change.
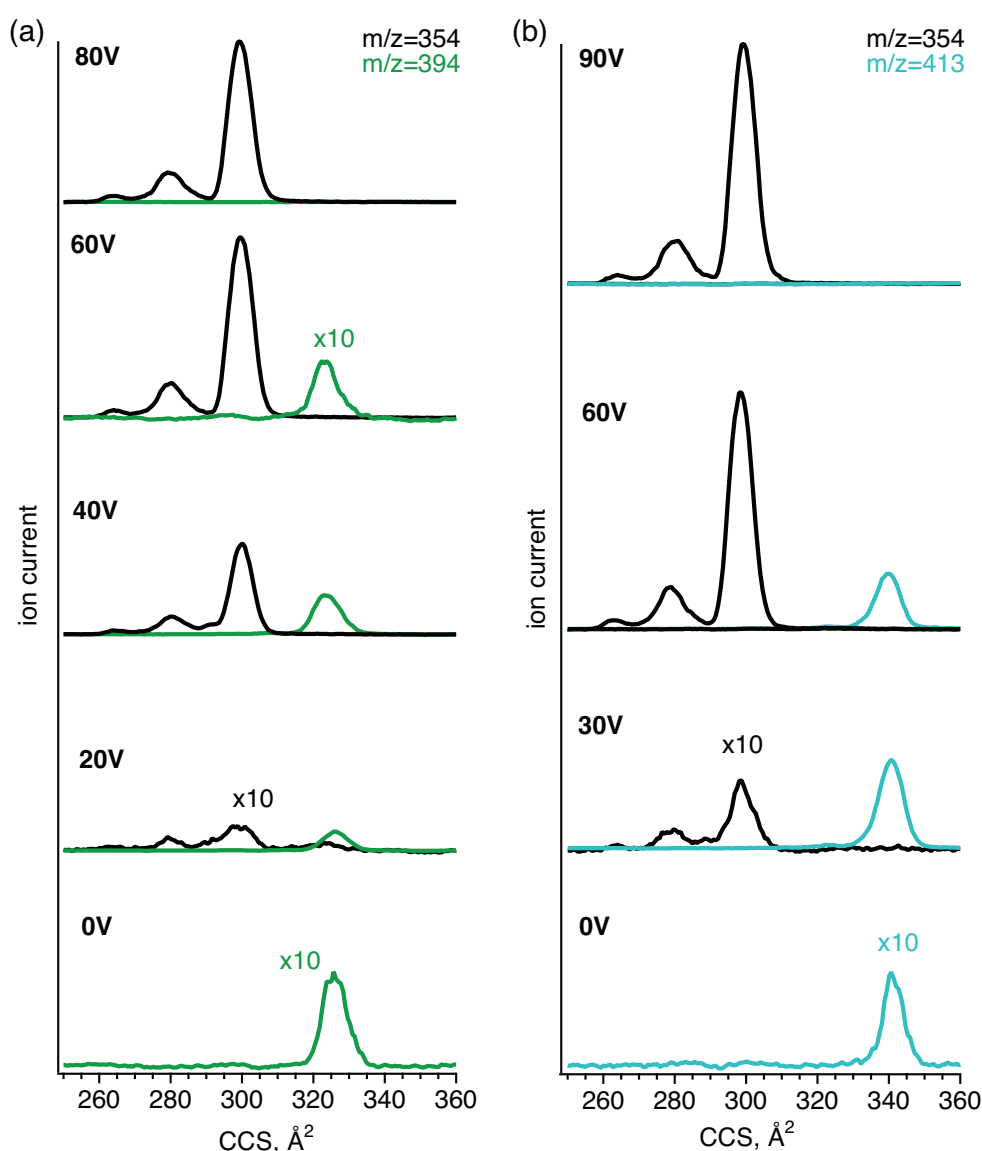
Figure 4-3. Collisional activation of pre-selected peaks with high CCSs in the IM-MS map. (a) The peak at 323 Å^2 with m/z=394 is selected and activated with 20 V, 40 V, 60 V and 80 V. The products of the activation are monitored on two different mass-to-charge ratios: on the original m/z=394, showed in green, and on the mass to charge ratio of bare triply protonated BK, showed in black. (b) The peak at 341 Å^2 with m/z=413 is selected and activated with 30 V, 60 V and 90 V. The products of the activation are monitored on two mass-to-charge ratios: on the original m/z=413 (blue trace) and on the mass to charge ratio of bare triply protonated BK (black trace). Some of the intensities are multiplied by a factor of 10, which is labeled with "x10".

4.3 FAIMS separation of BK^{3+}

Along with the IMS, we used FAIMS to separate conformational families of triply protonated BK. Figure 4-4(a) represents the CV distribution recorded by detecting all transmitted ions as a function of the compensation voltage. We found that introducing FAIMS as a pre-filter in front of the ion source of the CIS instrument does not significantly

deteriorate the total ion signal. The line shape of the CV distribution clearly suggests the presence of at least three distinct conformational families. While the features are not fully resolved, one can use photofragment spectroscopy to decompose the overall distribution into contributions from different conformational families.[360] To do so, we first set the CV at fixed values (indicated by colored arrows in Figure 4-4(a)), which only admits a subset of the parent ions into our cold ion trap. We then record UV photofragment spectra of these pre-selected ions, shown in Figure 4-5 for CV values of -7 V, -9 V and -11 V.
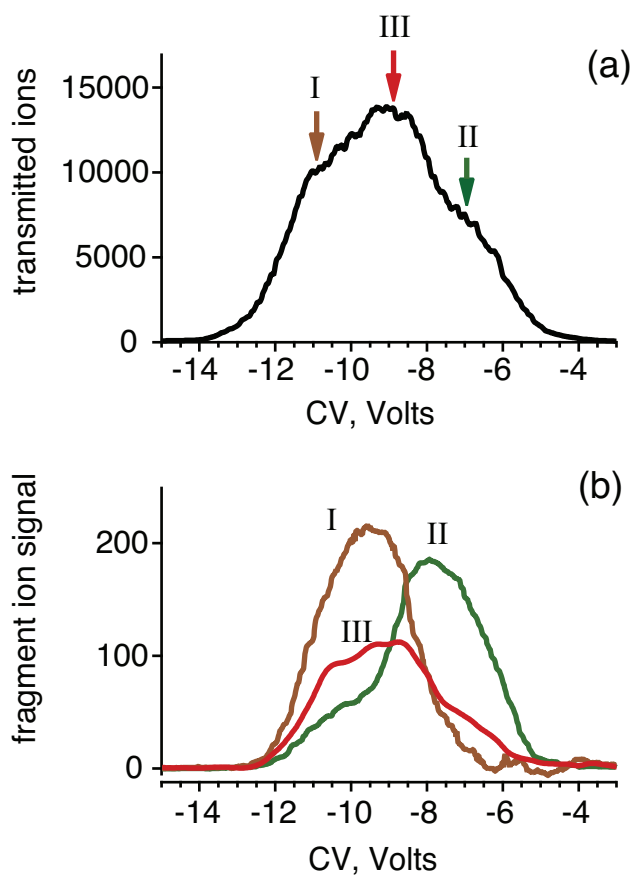
Figure 4-4. CV spectra of BK^{3+}; (a) detecting all transmitted ions of BK^{3+}, arrows show the CV values used to record the spectra in Fig. 2; (b) CV spectra recorded *via* UV transitions unique for conformational families I, II and III (arrows in Figure 4-5).

The UV spectra taken at the highest and lowest CV values, shown in Figure 4-5 (a) and (b), have no clear features in common, and these can be attributed to two different conformational families, which we call I and II. The lack of common spectral features indicates that these families are cleanly separated by FAIMS and that no isomerization between them occurs downstream of the separation step. The spectrum of Figure 4-5 (c),

which is taken at CV=-9V, contains peaks that belong to families I and II but also unique ones that we assign to a third conformational family (family III). To show more clearly the spectral features unique to family III, we subtract the spectra of families I and II with appropriate coefficients from the spectrum of Figure 4-5 (c). This difference spectrum is displayed in Figure 4-5 (d).
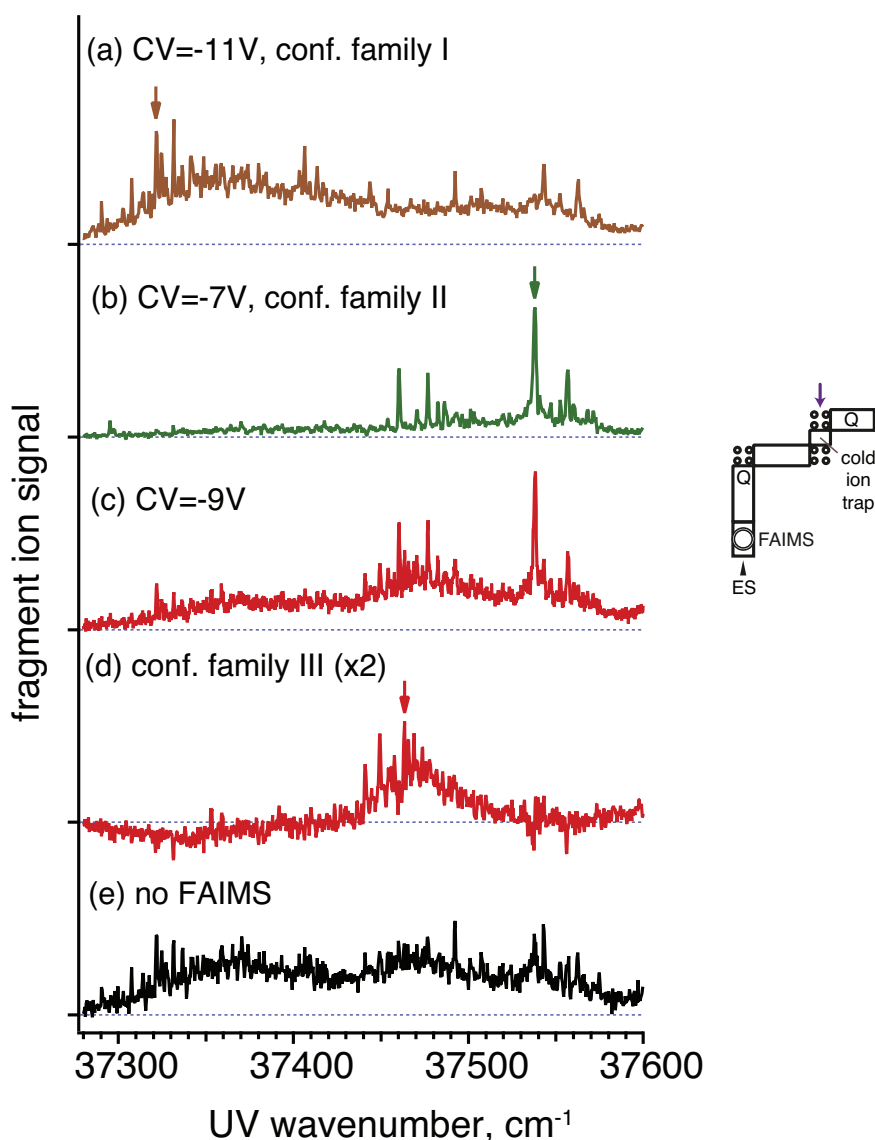
Figure 4-5. UV photofragmentation spectra of conformational families of BK^{3+}: (a) conformational family I, CV is set to -11V, (b) conformational family II, CV set to -7V, (c) a mixture of conformational families I, II and III, CV set to -9V, (d) conformational family III obtained by subtracting spectra of families I and II from spectrum c, (e) mixture of all conformational families, no FAIMS selection. Arrows show the UV transitions used to record conformer-selective CV spectra in Figure 4-4 (b).

With these UV spectra of distinct conformational families, we can set the UV laser on a unique spectral feature for each conformational family (37321.6 cm^{-1} for family I, 37538.0 cm^{-1} for family II, and 37463.6 cm^{-1} for family III), shown with arrows in Figure 4-5, and scan the CV. This produces a CV distribution for each conformational family, shown in Figure 4-4 (b). Collisional heating of the molecules during FAIMS separation also contributes to broadening of the peaks and produces shoulders on the CV-distributions of families II and III.

The CV distribution recorded on the UV transition unique for family III (red trace, Figure 4-4(b)) is broad and basically reproduces the total CV distribution of BK^{3+}. Together with other evidence showed below, it suggests that the conformational family III is mainly produced in the gas phase from families I and II, which can be found in the conformational distribution directly after the ESI process.

Nevertheless, this type of decomposition guides us in pre-selecting different conformational families and in understanding possible interconversion paths between them.

4.4 Spectroscopic evidence for kinetic trapping

The UV photofragment spectrum of the mixture of conformers produced by ESI without the FAIMS separation stage is shown in Figure 4-5(e). This mixture should correspond to the same mixture of conformational families observed using conventional ion mobility under same solution conditions. In general, we cannot reliably extract the relative abundances of the conformational families from UV photofragment spectra, because the intensities of the peaks are determined not only by the number of parent molecules, but also by the absorption cross-section and photofragmentation yield. However, since these parameters remain constant, we can interpret changes in relative intensities of the peaks as changes in number of ions adopting different conformations. We can thus use these spectra to address the question of whether these families are kinetically trapped, high-energy structures or simply represent the most stable gas-phase structures.

To answer this question we compared UV spectra of BK^{+3} with minimal and maximal degrees of collisional activation, using the RF amplitude of our ion funnel to control this. It is possible to change this amplitude from 7 V to 90 V before the BK^{+3} ions begin to fragment. We first select one conformational family at a time in FAIMS by fixing the CV on a given value, and then gradually increase the RF amplitude. After collisional activation close to the funnel exit, the ions are thermalized at room temperature in the hexapole pre-trap for 70

ms. The ions are subsequently mass-selected and transmitted to the cold ion trap, where they are cooled before a UV photofragment spectrum is recorded.

When conformational family I is selected (Figure 4-6 (a), CV=-11 V), we initially observe intense absorption bands at 37300-37350 cm^{-1}, which disappear upon increasing collisional activation. At the same time the bands corresponding to families II and III grow in, and at the highest level of activation the distribution of conformers consists exclusively of these two families. This means that conformational family I is kinetically trapped, and its population in the gas-phase after annealing is negligible. When we set the CV at -7 V to select family II and increase the degree of collisional activation (Figure 4-6 (b)), the spectra indicate that some of the molecules isomerize to family III. The spectrum at high collisional activation is practically identical to the one obtained at CV=-11 V, implying that we reach the same quasi-equilibrium gas-phase conformer distribution independent of the starting conformation. We do not observe the formation of family I, which lies higher in energy than families II and III in the gas phase. Finally, when the CV is fixed at -9 V, the distribution of conformers at an RF amplitude of 90 V is essentially the same as when families I and III were pre-selected, with no new conformations formed (Figure 4-6 (c)).

Directly after the ESI process all three conformational families are observed, while after annealing only families II and III are present. Looking only at the distribution of conformations in the gas phase, it is not possible to determine which environment is responsible for stabilizing family I. It may reflect the bulk solution-phase structure, but it could also reflect the specific environment of the electrosprayed droplets. The spectrum of Figure 4-6 (c) shows that most of the population in conformational family III comes from collisional activation of other families in the gas phase: it is less present than family II even when the CV is set to the maximum of its CV distribution (-9 V), and upon collisional activation its relative intensity increases. Family I is clearly the most important for addressing biologically relevant questions about structure and function of the peptide in solution.
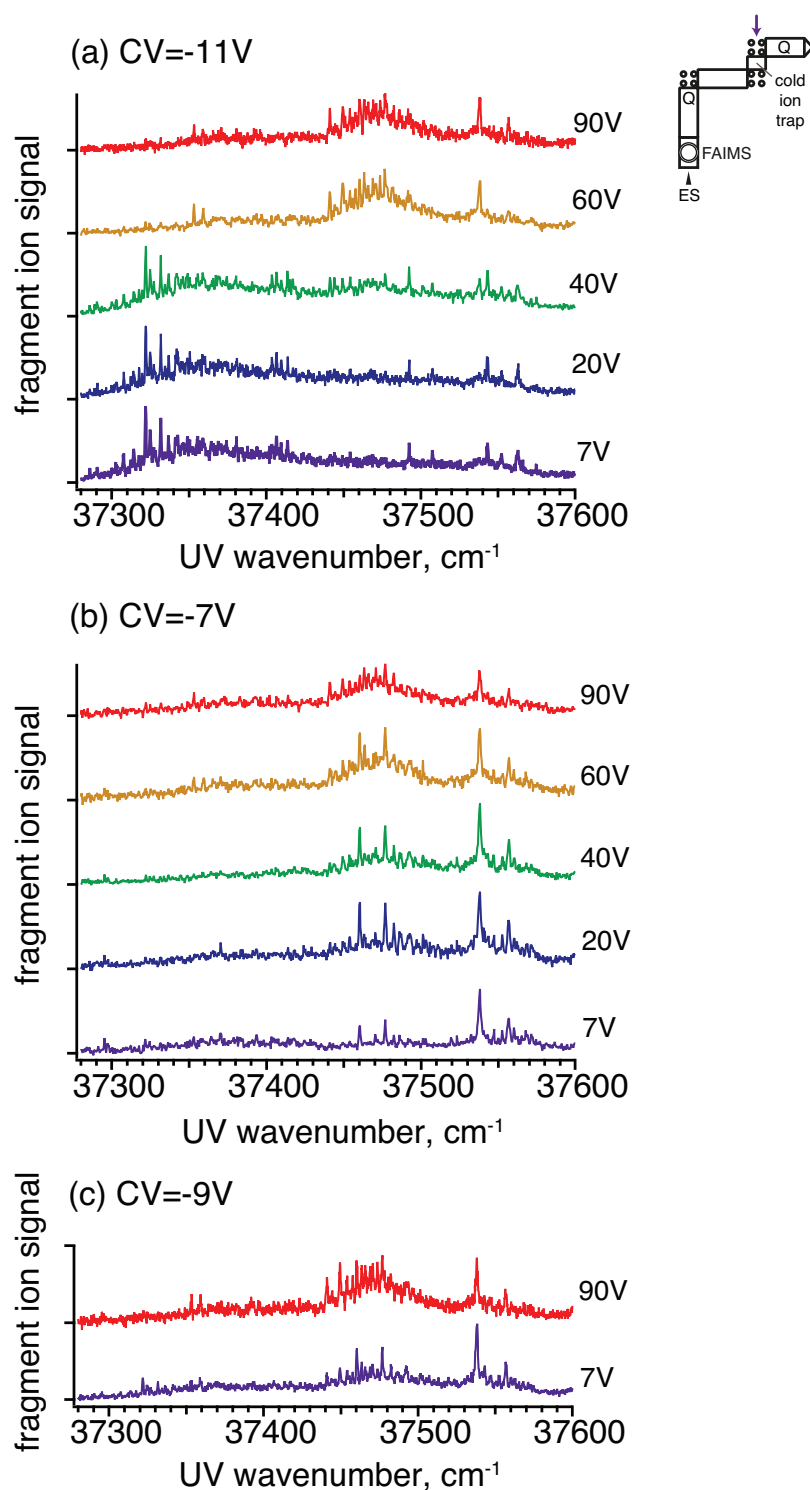
Figure 4-6. Collisional activation of each conformational family of BK^{3+}, preselected by FAIMS. The numbers on the right correspond to the amplitude of the RF voltage on the ion funnel in the source of the spectrometer. (a) Conformational family I converts to a mixture of II and III, (b) conformational family II converts partially to III, (c) the initially selected mixture of families I, II and III converts to the same distribution of stable gas-phase conformers as in a and b.

The behavior of three peaks separated by FAIMS is similar to that of the three conformational families observed using drift-tube IMS.[139] Shvartsburg *et al.* showed previously that there is some degree of correlation between these two separation techniques.[358] We suggest that what we designate as conformational family I corresponds to what was identified as conformer A in the ion-mobility studies of Clemmer and coworkers (labeled "A" in Figure 4-2 (a)).[139] The relative abundance of A in their gas-phase quasi-equilibrium distribution is around 2% of the total number of ions,[139] which is below the signal-to-noise ratio in our experiments, especially taking into account the quite broad UV photofragmentation spectrum of family I.

4.5 Spectroscopic characterization of the conformational families

4.5.1 *Vibrational spectroscopy*

The advantage of combining cold-ion spectroscopy with ion mobility is that instead of obtaining a single number, the CCS, with which to characterize a conformation, we can obtain a conformer-specific infrared spectrum, which provides a stringent test of calculated structures. Even without calculations, such spectra can provide qualitative information on structural features. Moreover, we can look in more detail at a conformational family separated by IMS or FAIMS, see how many different individual conformations it contains, and identify structural differences between them. We used double resonance IR-UV spectroscopy to obtain IR spectra of each conformational family in the region of NH and OH stretches. Applied to conformational family II, it reveals two distinguishable IR spectra, shown in Figure 4-7, and thus two different conformers present, which we call α and β. These two conformers combine to give rise to the UV spectrum of family II.

The gain at 3505 cm^{-1} in the spectrum of conformer α is produced by broadening of the UV spectrum of β conformer, when the IR laser hits a transition, which belongs to this conformer.   This effect in general limits our ability to   obtain   conformer-selective vibrational spectra of the minor conformers: instead of the dips, corresponding to the vibrations of this conformer, we mainly observe gain resulting from the broadening of the UV transitions of the major conformers.   This effect can be decreased *via* physical separation of the structures entering the cold trap: in this way the major conformer is no longer present in the trap and the spectroscopic study of the minor one can be performed with high signal-to-noise ratio. This argument supports the utility of separation techniques,

such as FAIMS, in spite of the fact that the physical foundation of FAIMS separation is not well understood, and the CV values cannot be directly converted to structural parameters.
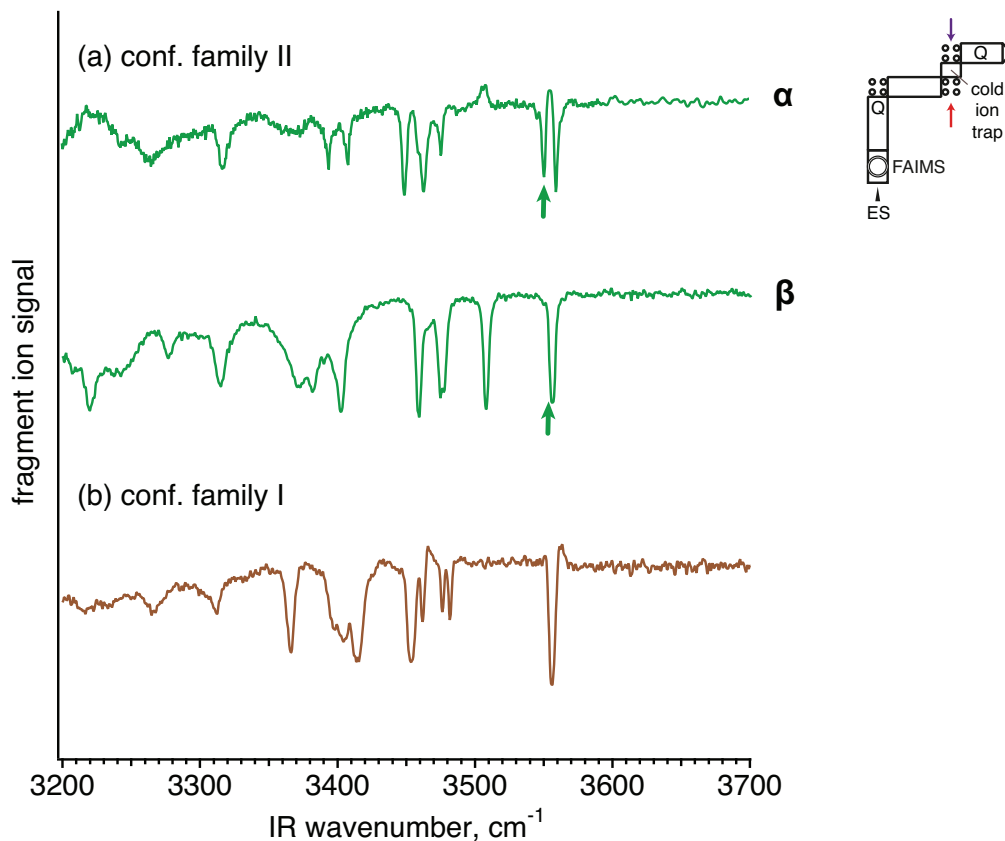
Figure 4-7. (a) Vibrational spectra of two major conformers of family II: α – UV laser is set on 37460.4 cm^{-1}, β – on 37538 cm^{-1}. The spectra recorded on all the other major transitions in the UV photofragmentation spectra are the same as for α or β. The arrows indicate the transitions used to separate the electronic spectra in Figure 4-8. (b) Vibrational spectrum of one of the conformers that belong to kinetically trapped conf. family I, recorded via setting the UV laser on the strongest transition of family I, 37321.8 cm^{-1}.

The UV photofragmentation spectrum of the kinetically trapped conformational family I is shifted significantly toward lower frequencies, indicating a strong interaction between the chromophoress and other parts of the molecule. This is consistent with this structure being the most compact.[139] The barrier holding this metastable structure in place is likely to arise from the competition between this tight hydrogen bonding and the mutual coulomb repulsion between the three charged groups. Upon collisional activation the structure seems to open up, which is confirmed by the larger CCS measured by Clemmer et al.,[139] the shift of the UV photofragmentation spectrum towards free phenylalanine (see Figure 4-6 (a)), and the blue shift of some of the NH stretch bands.

4.5.2   Conformer-selective UV photofragmentation spectra

We can also use the double-resonance nature of our technique to disentangle the UV spectra of the two conformers of the conformational family II. For each conformer we choose a strong band in the IR spectrum that does not appear in the spectrum of the other conformer: 3551 cm^{-1} for conformer α and 3556.3 cm^{-1} for conformer β (see Figure 4-7). The IR laser is then fixed on this wavenumber and the UV laser is scanned. All peaks in the UV photofragmentation spectrum that belong to the IR-labeled conformer will be depleted. By subtracting the UV spectra without and with IR laser we get the UV spectrum of each separate conformer of family II, shown in Figure 4-8.
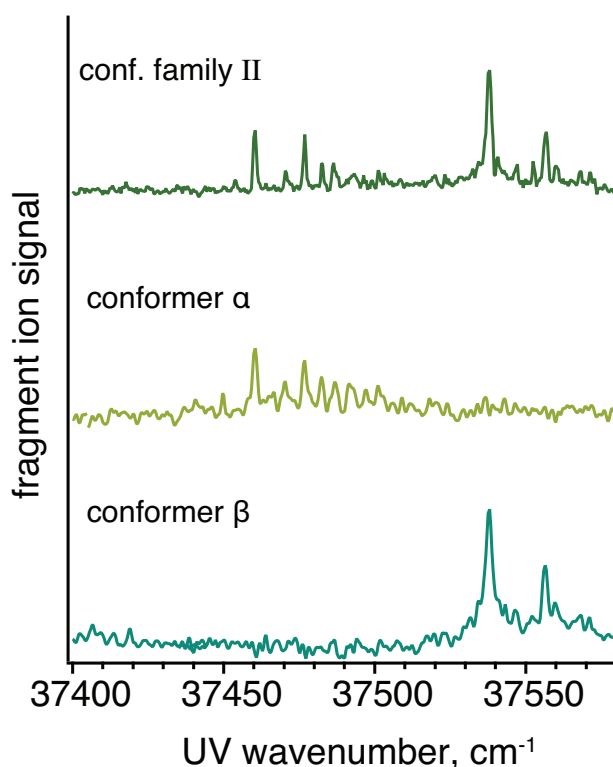
Figure 4-8. Decomposition of the UV photofragmentation spectrum of family II into a sum of spectra of two conformers, α and β, using IR-UV depletion spectra. Details of the procedure are discussed in the text.

These separate UV photofragment spectra provide some qualitative information about the structure of the conformers giving rise to them.   There are two phenylalanine chromophores in BK^{3+}, in the fifth and eighth positions in the sequence, and interaction between them could cause exciton splitting of the lines.[361] We do not observe it for either conformer of family II, suggesting that the side chains of phenylalanines are weakly interacting, either due to a large distance between them or their relative orientation. This

105

provides some guidance for structural search. Moreover, the band origin of conformer β is close to that of bare phenylalanine in the gas phase,[362] which means that the chromophores are free from any strong interactions. The band origin of conformer α is significantly red-shifted compared to bare phenylalanine, suggesting a strong cation-π or hydrogen-bonding interaction. Although they belong to the same conformational family, the structures of conformers α and β might be significantly different. The relative intensity of conformer α grows with respect to that of conformer β upon annealing (see Figure 4-6 (b)), which suggests that the former lies slightly lower in energy. This information could help when searching conformational space to identify the peptide structures of a given conformational family.

4.5.3 Assignment of the vibrational spectra

The task of assigning the vibrational transitions in the IR spectra of BK^{3+} is challenging due to the high number of oscillators in the molecule (Figure 4-9). In total there are 18 NH stretch vibrations and 2 OH stretches in triply protonated BK. In this work we consider only a part of the spectrum (3200-3700 cm^{-1}), which does not necessarily include all the bands, as some of the functional groups might be strongly hydrogen bonded and consequently their frequencies might be shifted below 3200 cm^{-1}.
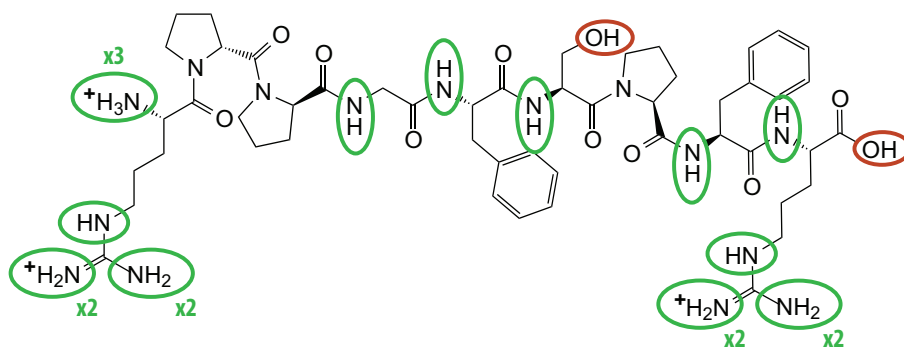
Figure 4-9. Chemical structure of triply protonated BK. Eighteen NH oscillators are labeled with green circles, two OH oscillators - with red.

Certain structural information can be obtained by qualitative analysis of the recorded IR spectra (see Figure 4-7). For example, the free OH of serine appears at ~3666 cm^{-1}, while none of the spectra of BK^{3+} we obtained have this spectroscopic feature.[227] It means that in all conformations the serine side chain is engaged in a hydrogen bond. There is at

least one band at ~3570 cm^{-1} in each conformer, which might correspond either to the carboxylic acid OH or to the arginine side chain.[344]

In order to assign the NH vibrations in the molecule, we synthetized several isotopically labeled compounds shown in Table 4-3. Each of them produces the same vibrational spectra as the non-labeled BK, except that a certain number of bands (the last column in Table 4-3), which correspond to the labeled oscillators, red shift by 8 cm^{-1} or less.

Table 4-3. Isotopically labeled compounds of BK. The "+" symbol corresponds to all ^{14}N in the amino acid being substituted by ^{15}N, the symbol "-" – to no labeling.

Name	^{1}Arg	Gly	^{5}Phe	Ser	Maximum number of shifting bands
L1	-	+	+	-	2
L2	-	+	-	+	2
L3	+	-	-	-	8

The spectra of compounds L1 and L2 combined allow us to unambiguously identify the NHs of glycine, phenylalanine and serine, as illustrated in Figure 4-10 for the case of conformer β. The band at 3476 cm^{-1} shifts in the spectra of both compounds L1 and L2 compared to non-labeled BK, which means that it belongs to the glycine NH stretch. We further deduce that the band at 3315 cm^{-1} is produced by the NH stretch of serine, while the band at 3457 cm^{-1} belongs to the ^{5}Phe NH. Note that the band at 3457 cm^{-1} does not shift entirely, but rather splits, revealing that two vibrational bands overlap. This is expected for a molecule of this size, as many similar oscillators are concentrated in a small spectral range.
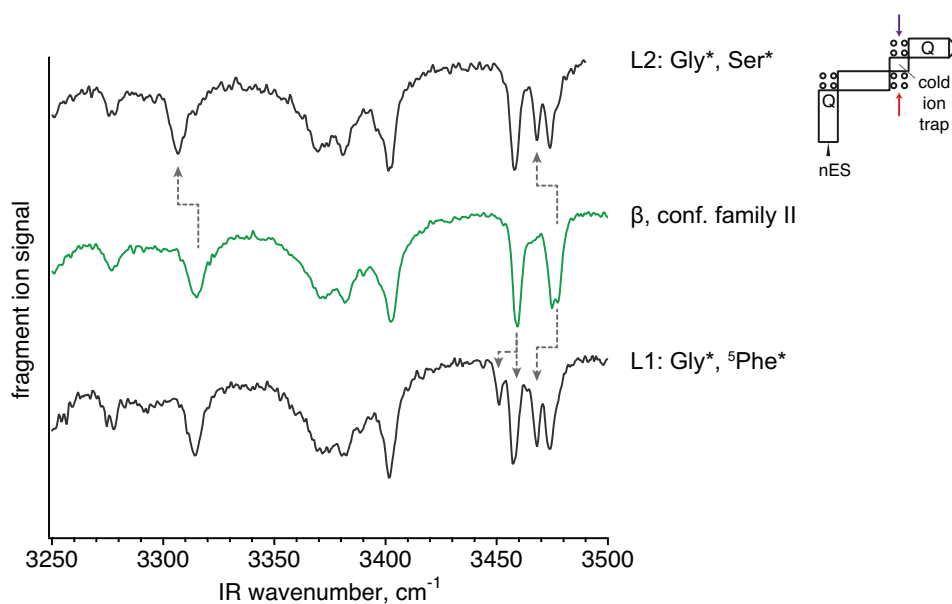
Figure 4-10. Shifts in the vibrational spectrum of conformer β of conformational family II induced by isotopic labeling. The labels on the right side indicate which amino acids in the sequence are isotopically labeled.

Isotopic labeling of all nitrogen atoms in the arginine in the first position potentially can lead to a shift of 8 vibrational bands. However, some NH stretches are strongly hydrogen bonded and thus either shifted below 3200 cm^{-1} or broadened, which makes a shift of 8 cm^{-1} difficult to detect. Figure 4-11 compares the spectrum of the most abundant conformer of conformational family I with an analog in which the first arginine is isotopically labeled. We observe 5 vibrational bands that either shift or split, indicating that they correspond to the side chain of arginine or to the protonated N-terminus. The band at 3554 cm^{-1} corresponds to the free asymmetric stretch of the η-NH$_2$ group, as it is the only NH-stretch that can be found at such a high frequency. Its symmetric counterpart corresponds to one of the bands at 3454 cm^{-1} or 3476 cm^{-1}. The free ε-NH should appear around 3465 cm^{-1} and might explain the second band out of the two listed above. The free N-terminus is expected to give rise to two bands at ∼3360cm^{-1} and 3300cm^{-1}; the former is not present in the spectrum.

Figure 4-11. Shifts in the vibrational spectrum of the most abundant conformer of conformational family I induced by isotopic labeling of the arginine in the first position.

We perform an analysis of the vibrational spectra of the isotopically labeled compounds L1-L3 for both conformers of conformational family II and for the most abundant conformer of family I. It allows us to identify the bands that belong to the labeled amino acids (Gly, 5Phe, Ser, 1Arg). In all measured conformers the bands shift by $\sim 8 cm^{-1}$, except for the 5Phe band in the spectrum of conformer α, which shifts by $\sim 4 cm^{-1}$. Since 8 cm^{-1} is the shift predicted for a free NH harmonic oscillator, a twice smaller shift indicates that the corresponding group is strongly coupled with another NH oscillator.

Up to now we have not identified the NH stretch of phenylalanine in position 8. Instead of using another isotopically labeled compound, we tentatively assign it using the vibrational spectra of the electronically excited S_1 state of the ions. An example is shown in Figure 4-12 for conformer β of conformational family II. The intense band at 3402 cm^{-1} shifts significantly to the red suggesting that it corresponds to a phenylalanine NH. Because we have identified the NH vibration of 5Phe using isotopic substitution and it lies at 3459 cm^{-1}, the band at 3402 cm^{-1} might belong to 8Phe, or, potentially, to another NH oscillator that is close in space to the phenyl ring.

Combination of isotopic labeling with the excited state spectra allows us to assign a large percentage of the vibrational bands (Figure 4-13). The transitions that are not assigned belong to either 9Arg NH stretches or the OH stretch modes (8 lines in total). This information is valuable for comparison between theory and experiment.
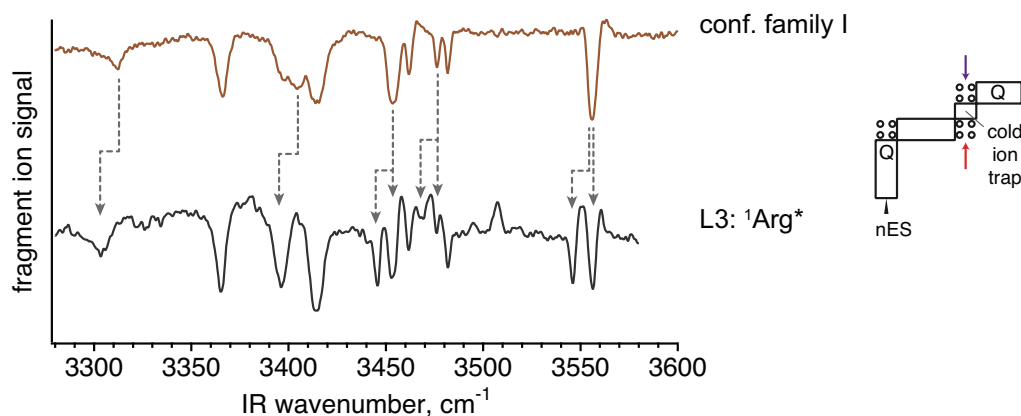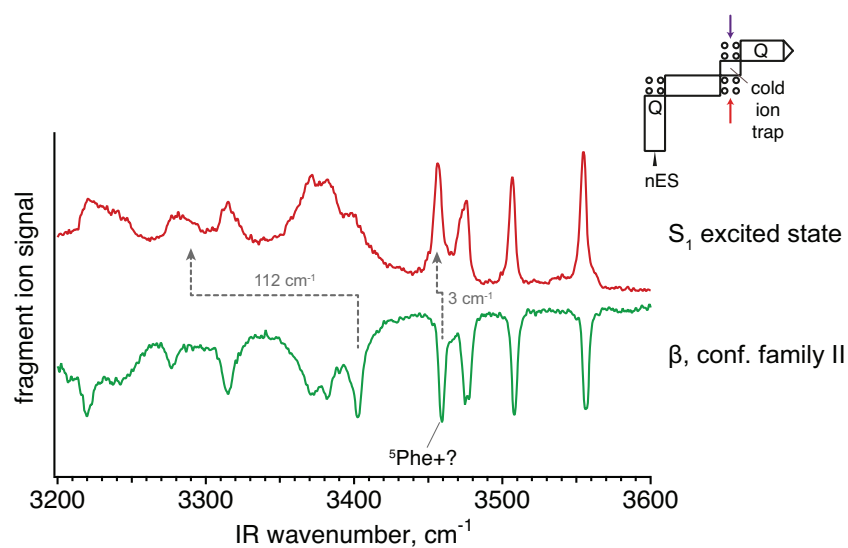
Figure 4-12. Shifts in the vibrational spectrum of the most abundant conformer of conformational family II induced by electronic excitation. The band corresponding either to phenylalanine in position 5 or to arginine in position 9 shifts by ~3 cm^{-1}. We tentatively assign the band that shifts by ~112 cm^{-1} to phenylalanine in position 8.
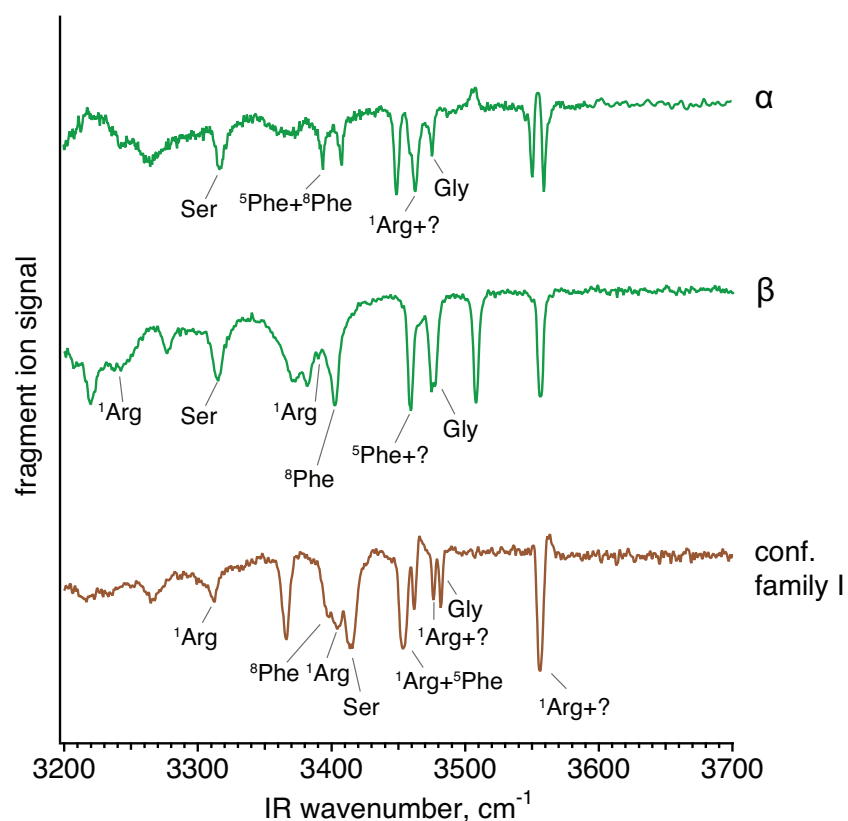
Figure 4-13. Vibrational spectra of BK^{3+} with the bands assigned using isotopic substitution and the excited state vibrational spectra.

4.5.4   Separation of two identical chromophores using isotopic labeling

The spectra of the electronically excite state of BK^{3+} raise a question: could we distinguish conformations using the fragmentation pattern upon electronic excitation? In principle, the UV photon can be absorbed by either of the two phenylalanines, producing a shift in the vibrational frequency of the corresponding NH and the loss of the respective side chain. Figure 4-12 suggests that in conformer β the UV light is mostly absorbed by the phenylalanine in position 8, since the frequency of its NH stretch shifts significantly. A slight shift is observed for the NH stretch of the phenylalanine in position 5 or for the arginine in position 9 (their bands overlap at 3459 cm^{-1}). The shift in the position of ^{9}Arg NH might be caused by electronic excitation of ^{8}Phe, supporting the statement that it is the only chromophore that absorbs the UV light in this conformation.

In the isotopically labeled compound L1, we have not only ^{14}N of ^{5}Phe substituted by ^{15}N, but also all carbon atoms of this amino acid substituted by ^{13}C. It means that the mass of the side chains of two phenylalanines is different and we can distinguish the fragments resulting from a side chain loss of ^{5}Phe from that resulting from a side chain loss of ^{8}Phe. A fragment mass spectrum of compound L1 recorded with the UV laser set at 37458.6 cm^{-1} is shown in Figure 4-14. We observe two well-resolved photofragments, which correspond to the side chain losses of two phenylalanines. Monitoring each of them separately, we record UV photofragmentation spectra, shown in Figure 4-15.
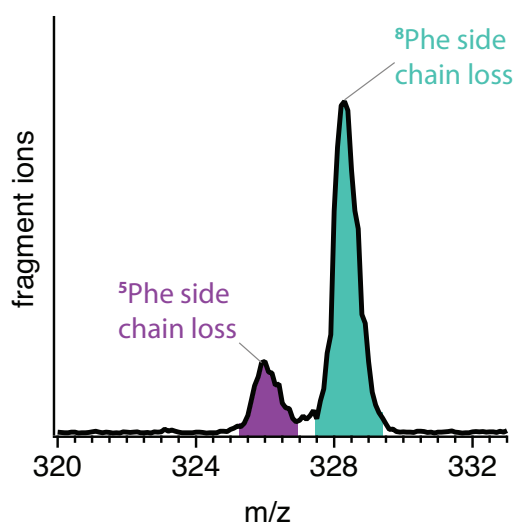
Figure 4-14. Mass spectrum of the laser-induced fragments of L1 compound with a "heavy" side chain of ^{5}Phe. The UV laser is set at 37458.6 cm^{-1} and it is followed by a CO_2 laser pulse. The CO_2 laser alone does not induce any observable dissociation at these masses.

111
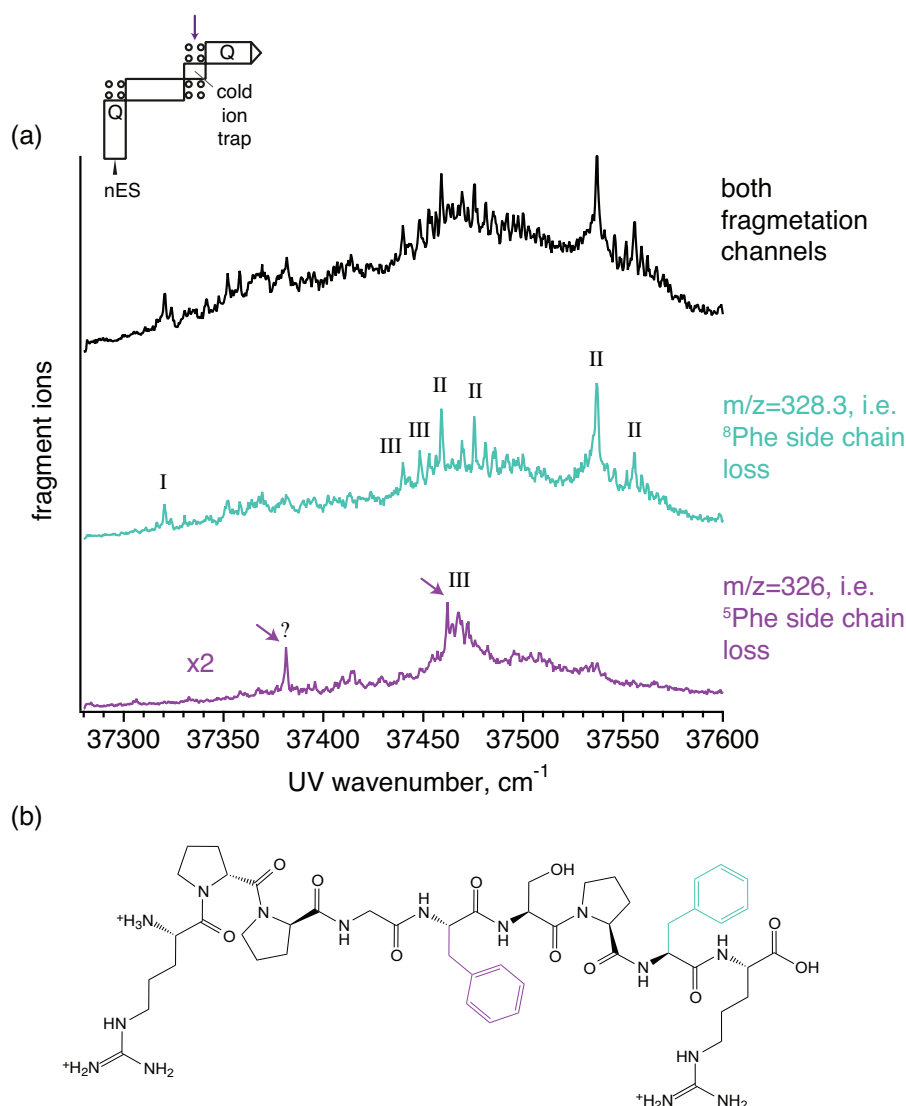
Figure 4-15.   (a) Electronic spectra of BK^{3+} recorded using different fragmentation channels, i.e. monitoring different mass-to-charge ratios of the fragments. The numbers indicate which peaks correspond to which conformational family separated by FAIMS. The arrows indicate the transitions for which the vibrational spectra were recorded.   (b) Chemical structure of BK showing the side chains of two phenylalanines, which fragment upon UV excitation. The colors correspond to the spectra in figure (a).

In general, UV light absorption could be followed by (i) resonance energy transfer to another chromophore, (ii) intersystem crossing and non-statistical fragmentation, as in the case of an ion with a single chromophore (see section 2.1.2), (iii) internal conversion and statistical fragmentation (same as collision-induced dissociation) or collisional cooling by residual-gas, (iv) fluorescence. It is clear that the first UV photon is initially absorbed by one of the two phenylalanine chromophores. Four different scenarios are plausible for a given chromophore:

- If the phenyl rings interact and the energy is equally distributed between them, we would record the same ratio of the photofragments resulting from [5]Phe and [8]Phe at any wavelength.

- At another extreme, if the chromophores do not interact, we would observe two sets of bands in the UV spectrum of a given conformer, each corresponding to absorption and subsequent fragmentation of one of the chromophores.

- An intermediate situation is possible if there is a limited energy transfer along with competing degradation channels.

- One could imagine that due to geometrical constraints only one phenylalanine side chain can fragment, whether it absorbs light itself or acts as an acceptor in the resonance energy transfer process.

The UV photofragmentation spectra are produced by a mixture of conformers, each of which can follow different scenarios presented above. The spectrum recorded monitoring the [8]Phe side chain loss contains all conformers that were studied previously, namely, the most abundant representatives of families I and II. In contrast, the spectrum recorded using [5]Phe side chain loss is composed of a part of the family III and of an intense line at 37381.4 cm^{-1}, which we could not attribute to any conformational family, as it was not pronounced in any of their UV spectra. The spectra in Figure 4-15 do not share any intense bands, which suggests that the chromophores do not interact in any of the molecular geometries. This is consistent with the UV absorption bands being close to that of bare phenylalanine.

We conclude that the most probable scenario is that each conformation preferentially fragments through only one of the side chain loss channels. Full investigation of the energy transfer process and the fragmentation pathways is beyond the scope of this work. Instead, we use the clear difference between the UV spectra recorded on different mass-to-charge ratios as an additional tool for conformer separation.

4.5.5   Vibrational spectrum of conformational family III

As stated in section 4.5.1, the case of triply protonated BK illustrates the need for several orthogonal methods of conformer separation in large systems. The most intense characteristic band of the conformational family III is located at 37461.8 cm^{-1}, and it is separated from an intense band of conformer α by only 1.6 cm^{-1}. When the IR laser hits a

vibrational transition of a particular conformer, the electronic spectrum of this conformer broadens, which results in a depletion of the photofragmentation signal on the UV transition of this conformer and a gain on the surrounding UV wavelengths, as shown schematically in Figure 4-16(a).[216]   As a result, when the UV laser is parked on the transition corresponding to family III and the IR laser comes in resonance with a vibrational transition of conformer α, a significant gain is observed, which obscures the depletion coming from the conformer of family III (Figure 4-16(b)). This obstacle can be overcome by subtraction from the recorded spectrum of the contribution of conformer α, but this procedure results in a very poor signal-to-noise ratio. This issue, in turn, can be resolved by significantly increasing the acquisition time. In Figure 4-17(a) we present a vibrational spectrum of one of the conformers of the conformational family III recorded this way.
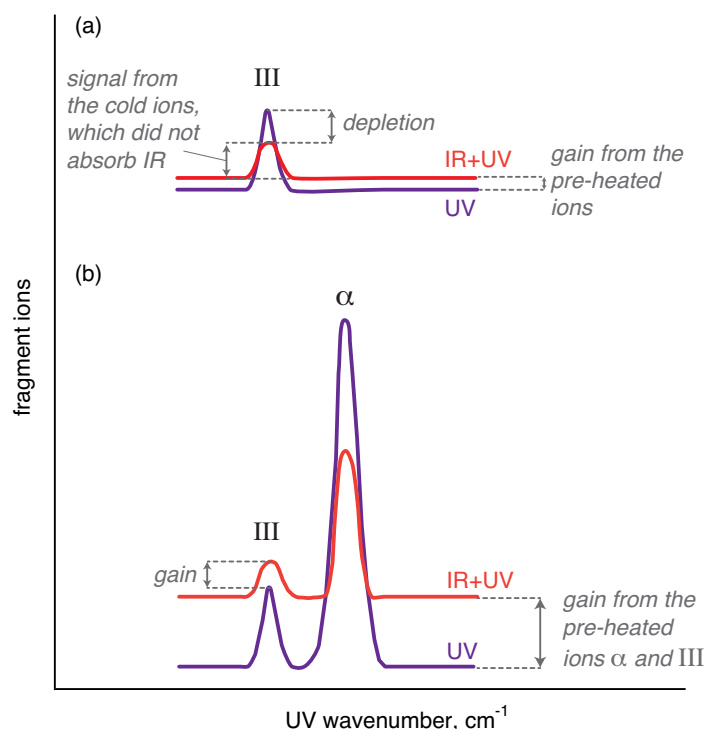
Figure 4-16. Schematic representation of the change in the electronic spectrum upon heating the ions with an IR pulse prior to UV excitation. (a) An ideal case: the UV band of the conformer of interest (III) is separated from other peaks in the electronic spectrum. The red trace shows the electronic spectrum of the ions after absorption of an IR pulse, which is in resonance with a vibration of conformer III. Scanning the IR laser and having the UV laser set on the transition of conformer III, the typical depletion spectrum is recorded. (b) The purple trace shows an electronic spectrum, in which apart from conformer III, a band of an abundant conformer α is present. If the vibrational transitions of conformers α and III overlap, the contribution of conformer α to the broadening of the electronic spectrum might be larger than the depletion of the band of conformer III (red trace). Scanning the IR laser and having the UV laser set on the transition of conformer III, we record a sum of depletion produced by conformer III and gain resulting from conformer α.

114
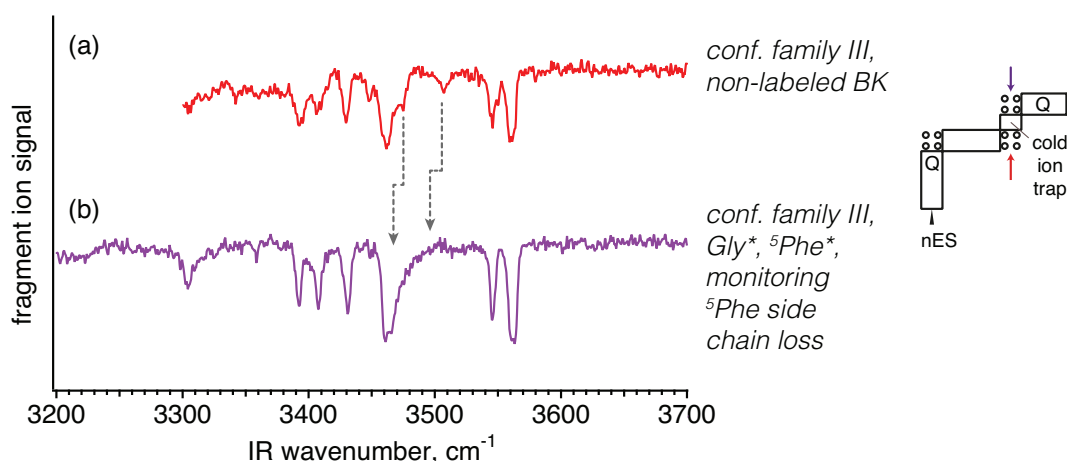
Figure 4-17. (a) Vibrational spectrum of the most abundant conformer of family III. The UV laser is parked at 37461.8 cm^{-1} and the spectrum is recorded monitoring a sum of both side chain loss fragments. The gain component resulting from conformers α and β is subtracted, which reduces the signal-to-noise ratio. (b) Same as (a), but recorded monitoring ^{5}Phe side chain loss signals using an isotopically labeled compound L1. The dashed lines indicate the shifts in the vibrational spectra induced by isotopic labeling.

On the other hand, when the conformers are spatially or spectroscopically separated, the spectrum of family III can be recorded directly. Unfortunately, using FAIMS we do not achieve full separation of families III and II (see Figure 4-6). Thus separation through the difference in the fragmentation patterns appears to be a useful tool, which facilitates double-resonance spectroscopy. Figure 4-17(b) shows the vibrational spectrum of the isotopically labeled compound L1, recorded monitoring m/z=326.0. No additional data treatment is required and the acquisition time is significantly shorter than that for Figure 4-17(a). Note that two bands in the IR spectrum are shifted compared to the non-labeled species, since the NH stretches of Gly and ^{5}Phe and isotopically labeled in the L1 compound.

4.6   Considerations of collisional activation and cooling

To put the conformer distributions that we measure into proper context, it is important to consider the collisional activation and cooling processes that we subject them to. In our CIS instrument, when no IM pre-filter is used, the ions are initially generated by electrospray at room temperature and pressure and injected into the spectrometer. After passing through a metal capillary, they enter the RF ion funnel, where we can activate them (primarily near the exit) by an increase in RF amplitude. After exiting the ion funnel, ions are collisionally cooled in two steps: first during a 70 ms pre-trapping period in a hexapole ion trap at room temperature, and subsequently in the cold ion trap, where the ions collide with cold helium at ~4 K. The final distribution of conformers that we observe in the

spectroscopic experiments depends not only on the initial distribution of conformations produced by electrospray, but also the level of collisional activation in the ion funnel and the rate of collisional cooling. If the ions are cooled significantly faster than the interconversion rate between different conformers, one obtains a "snapshot" of all the conformations present in the ion cloud before cooling. In this case the spectroscopically observed distribution of conformers reflects the equilibrium distribution corresponding to the initial temperature of the ions. If the cooling process is significantly slower than the isomerization, cooling can be considered as adiabatic process: equilibrium would be established at every given moment, and one would always find a single conformer family, as the system would have enough time to interconvert to the lowest-energy conformer. This is obviously not the case in the present experiments since we observe multiple conformations, which means that either cooling in the hexapole pre-trap or in the cold octopole trap happens relatively fast. The cooling rate in the hexapole should be lower than in IMS experiments due to the lower pressure: 7e-3 mbar compared to several mbar. This can decrease the percentage of higher-energy (i.e., kinetically trapped) species in the distributions measured in our CIS spectrometer, since they correspond to lower effective temperatures.

4.7 Considerations for benchmarking quantum chemical calculations

The interplay between quantum chemistry calculations and measured vibrational spectra of gas-phase biological molecules is bi-directional. On the one hand, the interpretation of measured spectra requires computed structures and spectra. The normal procedure is to calculate the lowest energy structures and their vibrational spectra, and then compare them to experiment. The experiments are used to select which of the computed structures is the "correct" one, and the calculations are used to interpret the vibrational spectra, particularly in cases where the measured spectra are not assigned. On the other hand, spectroscopic signatures of peptides are frequently used to benchmark quantum chemical calculations.[213,214,363-366] The comparison between calculated and measured IR spectra serve to evaluate the performance of new methods or functionals. For this theory-experiment interplay to work, one needs to know whether measured IR spectra should be associated with the lowest-energy conformations for which the quantum chemical calculations typically search. If the conformations measured in the gas phase do not correspond to the lowest in energy but are kinetically trapped metastable structures, conformational searches may never find them. Moreover, if the accuracy of computational

methods is judged by a comparison of the measured spectrum of a metastable species with the calculated spectrum for what is found to be the lowest energy structure, the benchmarking procedure may be completely misguided. From the experimental side, it is thus important to know when the species one measures may be kinetically trapped, so that computational procedures can be guided to look for higher energy structures. From the computational side, methods need to be implemented to predict how solution phase structures might be trapped during transfer to the gas phase, so that the proper regions of conformational spaces are explored. One of the possible approaches to meet these requirements is demonstrated in section 3.3.

Triply protonated bradykinin provides an example of how the interplay between experiment and theory could go wrong. Figure 4-18 represents the UV photofragmentation spectrum of BK^{3+}, which we originally obtained in our cold-ion spectrometer without any FAIMS pre-selection of conformational families. This spectrum corresponds neither to the stable gas-phase structures exclusively nor to the distribution directly after the electrospray process – it is dominated by conformational family II. Quantum-chemical calculations may thus be misguided by these data. In this case, the experimentally observed distribution is practically missing conformational family III, which is formed upon annealing and thus likely to be the lowest-energy in the gas phase. Comparing results of quantum-chemical calculations with the vibrational spectra of family II would explore only a part of all the relevant conformational space of BK^{+3} in the gas phase.
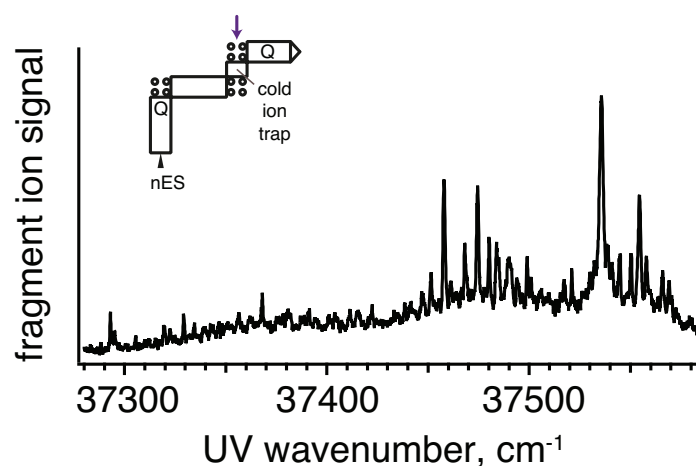
Figure 4-18. UV photofragmentation spectrum of BK^{3+}, obtained in conditions optimized for maximum of ion signal. No FAIMS separation was applied.

As spectroscopic techniques are pushed to increasingly larger peptides, one would like to measure spectroscopic signatures of kinetically trapped structures, as they may be the most relevant for understanding the function of peptides in solution. Quantum chemical calculations should then be adapted to search for geometry of ions based not only on energy threshold but also on experimental values such as CCS and IR spectra. Finding kinetically trapped conformations remains challenging, as they may lie relatively high in energy and are separated from stable gas-phase structures by substantial barriers. However, knowledge about their structure is essential to link gas-phase measurements with the structure of biological molecules in solution.

Detailed understanding of the cooling mechanism is also essential for molecular modeling. Relative free energies of different conformers change with temperature,[364] and measured conformer distributions should be assigned to a particular temperature in order for the comparison between theory and experiment to be correct. If the cooling process in the cold ion trap is assumed to be infinitely fast compared to isomerization, the measured distribution reflects distribution of conformers at room temperature. This effect is also illustrated in section 3.3.2 for the *trans-trans* conformations of BK[1-5]$^{2+}$.

4.7.1   *Molecular dynamics to predict the proline conformation*

In Chapter 3 we showed that the proline configuration alone does not predict the CCS of the bradykinin 1-5 fragment: for any possible proline conformation both compact and extended conformations are achievable. Here we test this statement for a larger system containing 9 amino acids. We performed replica exchange molecular dynamics simulations using the OPLS-AA force field and clustered the resulting structures. In this way we obtained 11,475 distinct conformations, which we then separated into 8 classes according to the proline conformation (CCC, CCT, TTC etc.) and computed their CCSs using the PA method. We then plotted the force field energy *vs* CCS in every category (Figure 4-19). For the computed CCSs a scaling factor of 1.014 was used to bring them close to the results of the TM method (in the same way as in section 3.3).
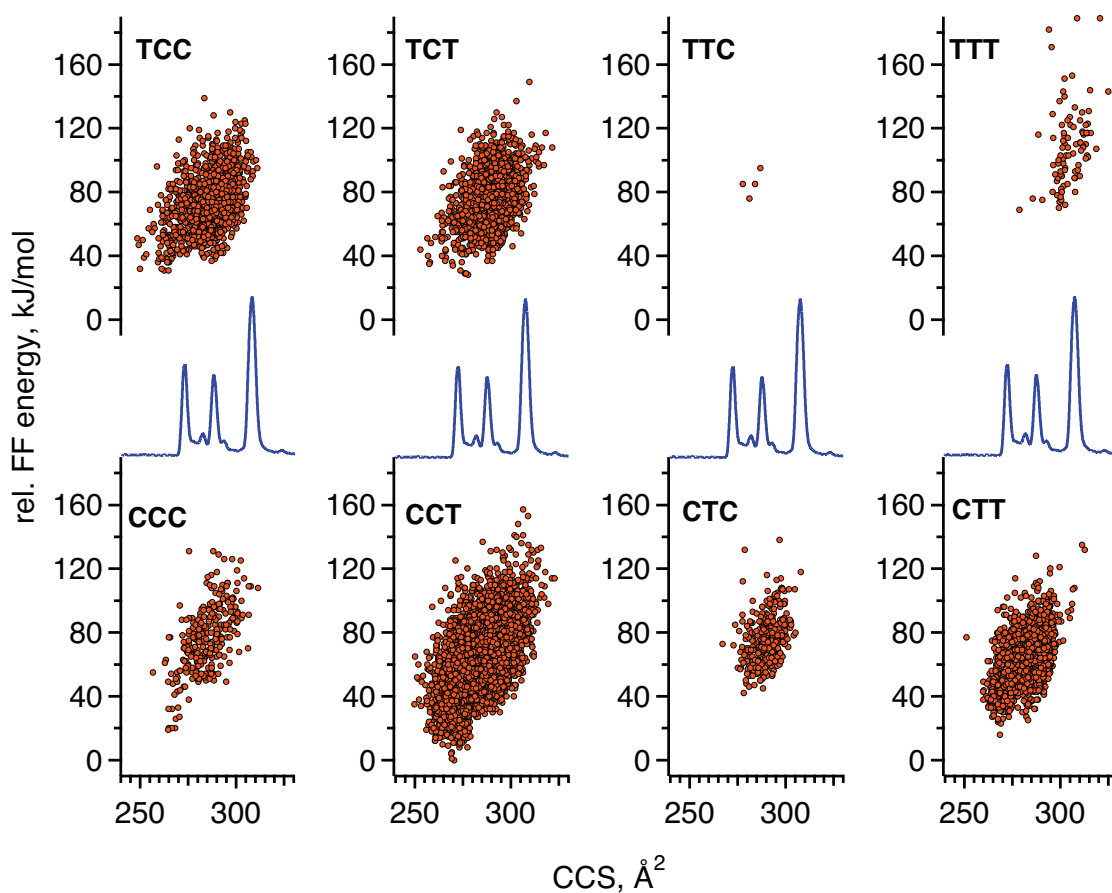
Figure 4-19. Relative force field energy of BK^{3+} in the gas phase as a function of the CCS for all possible proline conformations. The energies are relative to the lowest-energy structure, which belongs to the CCT type. The experimental IM distribution in blue is shown for comparison.

The most populated backbone conformations such as CCT and TCT span across all peaks in the IM distribution. High barriers on the PES formed by proline *cis-trans* isomerization might form the distinct conformational families, but Figure 4-19 strongly suggests that the CCS itself does not correlate with the proline conformation.

The lowest-energy conformation belongs to the CCT type. This is not a definitive result, as we have observed previously that the relative energies might change dramatically when re-optimizing the geometries using the DFT. The CCT type was suggested to be a minor sub-population of peak C.[70] Surprisingly, we find only 4 representatives of the structures that are thought to comprise the main peak in the IM distribution, TTC.[173]

In order to evaluate how the CCS *vs* energy distributions change as the level of simulations increases, we optimized all geometries in TTT and CTT types with DFT using PBE functional and "light" convergence settings (Figure 4-20).
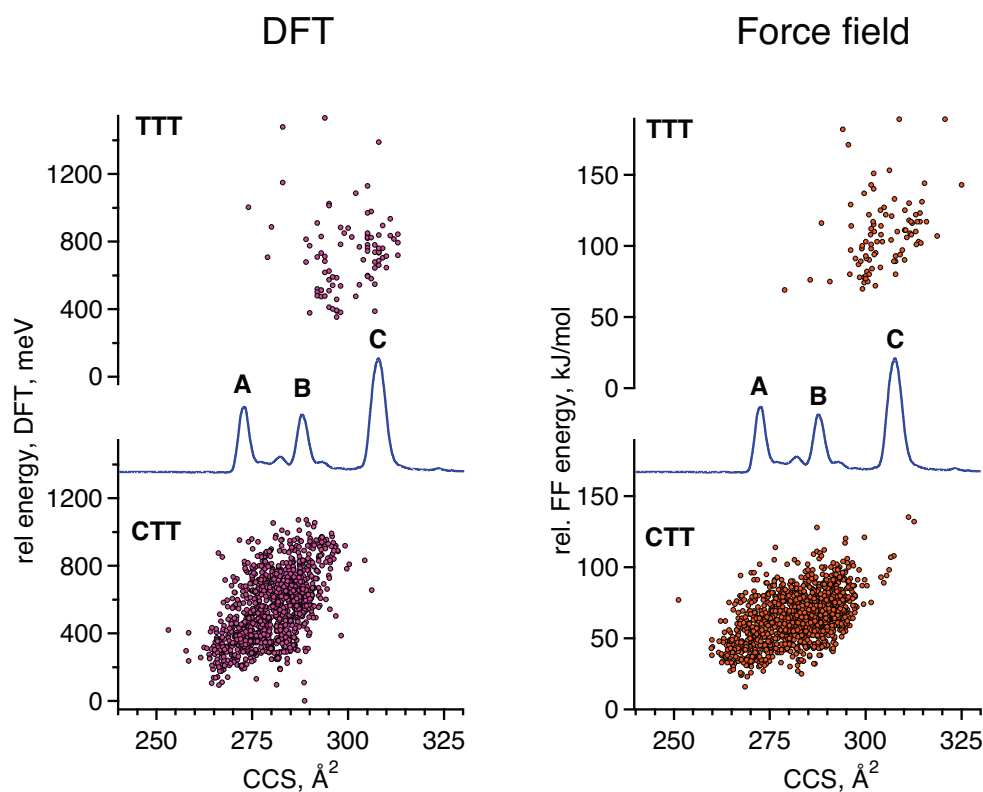
Figure 4-20. Comparison between the energy *vs* CCS distributions computed using DFT and force field calculations. Two conformational types TTT and CTT are presented. The DFT energies are computed relative to the lowest-energy structure of the CTT type. The experimental IM distribution in blue is shown for comparison.

The most populated solution structures, i.e. TTT, are clearly extended. They can have the CCSs corresponding to peaks B and C, but not A. The opposite is true for the CTT structures: they are compact and may be found within peaks A and B. Density functional theory optimization shifts the global minimum within this conformational family from peak A to peak B.

In general, Figure 4-20 proves that further optimization of all geometries with DFT is needed to identify the global minimum on the PES and to show that it falls within peak C in the IM distribution. Identification of the kinetically trapped structures requires comparison with the spectroscopic data, as the energy of these species is high by definition. The CCS alone cannot serve as a sufficient criterion, since all proline configurations except for TTT and TTC can form compact structures with the CCSs corresponding to peaks A and B (see Figure 4-19).

4.8 Conclusions

This work provides an overview of an ensemble of conformational families of BK^{3+} in the gas phase. Previously published IM studies allowed separation of conformational families and direct observation of their interconversion.[70,139] The double-resonance spectroscopic scheme allows further separation of conformational families into several conformers and facilitates their detailed characterization. We identified three conformational families separated by barriers on the potential energy surface, in agreement with previous ion mobility studies.[139] While ion mobility provides a collisional cross section for each conformational family, we obtained complementary data such as electronic and vibrational spectra, which can be used as benchmark for quantum chemical calculations of the structures of BK^{3+} in the gas phase. Knowing these structures with atomic precision would allow one to rationalize a vast amount of experimental data obtained for this peptide using different gas-phase techniques and compare them with solution-phase signatures of the conformational ensemble, bringing us closer to understanding the relationship between solution-phase and gas-phase structures of biomolecules.

From a biological perspective one is mostly interested in studying the structures that correspond to the solution phase, whether it is an aqueous environment or a biological membrane. Qualitatively our data confirm that the kinetically trapped structures of BK^{3+} in the gas phase are compact, in agreement with the statement that in an aqueous solution BK has a relatively compact globular structure as the charges on the side chains are screened by the solvent. Upon annealing in the gas phase the molecules interconvert from compact structures to a distribution of more extended ones, and might reach a conformational distribution close to that in a non-polar medium such as a biological membrane. It has been shown that BK interacting with one of its receptors is extended and has a well-defined S-shape with all prolyl-peptide bonds in the *trans* conformation.[185] Could the set of low-energy structures in the gas phase actually model the biologically active state of BK in non-aqueous media? The spectroscopic signatures of different conformers presented in this study could help us answer this question by direct comparison of the structures in the gas phase with previously characterized conformations in aqueous and non-polar solutions.

4.9 Additional experimental details

Bradykinin (RPPGFSPFR, acetate salt, 98%, Sigma Aldrich) was purchased and used without further purification. Peptide solutions were prepared in a 49:49:2 mixture of water: methanol : acetic acid with a bradykinin concentration of 50µM. In the 3+ protonation state, the extra protons are believed to reside on the two arginine side chains and on the N-terminus.[173] Isotopically ^{15}N-labeled bradykinin analogs were synthetized by Bachem AG and used without further purification.

Chapter 5.
Can proline to alanine mutations be used to assist structure determination?

In this chapter we assess the role of proline *cis/trans* isomerization in forming the conformational space of triply protonated bradykinin in the gas phase. To that end, we compare UV and IR spectroscopic signatures, as well as FAIMS and DT IM distributions, of native BK and two mutants, in which one or two prolines are substituted by alanine.

5.1 Introduction

Proline is the only canonical amino acid for which *cis* and *trans* isomers of the peptide bond are comparable in energy.[367] However, on average the percentage of *cis* peptide bonds preceding the proline residue is low (5.2% according to Weiss and co-workers[148]). To assess the role of these rare states, the possibility of isomerization can be eliminated by "locking" the Xxx-Pro bond to the *trans* isomer *via* substituting the proline with another amino acid.[150,173,368-372] In some cases such mutation does not affect the overall structure of the protein in solution,[151,371] while in other cases it leads to drastic changes in the macroscopic properties of the protein and in its NMR spectra.[150,151,372,373] For example, in antifreeze protein (PDB code P19614), which consists of 67 amino acids, the prolines at the C-terminus are not critical for the protein structure, but substitution of prolines 29 and 33 to alanine changes the antifreeze activity, alters the structure and decreases the solubility.[373] In a similar way, the peptide bond can be locked in the *cis* conformation by replacing proline with pseudo-

proline[374,375] or aminoisobutyric acid.[158] The opposite mutation of any other amino acid to proline (so-called proline screening) is used to characterize secondary structure of peptides and proteins by measuring how sensitive it is to such substitution.[370]

Mutational analysis, in conjunction with various methods of structural determination, shows that proline performs multiple functions within the protein sequence: it facilitates formation of certain types of β-turns,[161] acts as a molecular switch;[150-154] and its isomerization is often a rate-limiting step in protein folding.[151,155-157] Importantly, in some cases substituting prolines by their *cis* analogs leads to a molecule with very low biological activity,[158] despite the fact that in nature the *cis* form of proline is more evolutionary preserved than the trans form.[376] Apparently, it is not solely the ability to form a *cis* peptide bond, but the conformational heterogeneity induced by the proline residues that plays a key role in all functions listed above.[150,151,153]

In section 1.4 we summarized the effects of the presence of proline on protein structure both in solution and in the gas phase. In relation to ion mobility, it was suggested that proline is a primary reason for multiple resolved peaks in CCS distribution of short peptides.[165,173-175] In order to establish a correspondence between features in the IM distribution and conformational preferences of prolyl-peptide bonds, the same point mutations as in solution-phase studies are used.[70,142,172-174,377] For instance, three conformational families of triply protonated bradykinin in the gas phase were attributed to different *cis-trans* isomers using proline-to-alanine substitutions.[70,173] However, while using this technique in solution can be directly verified by comparing the NMR spectra of the mutant with those of the original sequence, such detailed study in the gas phase has not been performed yet.

Three significantly different conformational families of triply protonated bradykinin (BK[3+]) are observed in the gas phase both by DT-IMS[139] and FAIMS.[350] Each of them in turn consists of multiple molecular geometries.   How does the conformational space of BK[3+] change upon proline-to-alanine substitution? Intuitively, "locking" a peptide bond in the *trans* conformation restricts the conformational space of the molecule and a subset of the originally observed molecular geometries should be observed. In this chapter we test this hypothesis for two mutants of bradykinin using a combination of ion mobility spectrometry and double-resonance, cold-ion spectroscopy.

5.2 Non-substituted bradykinin

As discussed in section 4.3, the compensation voltage (CV) distribution of BK^{3+} produced in our FAIMS device shows the presence of several conformational families, though their peaks are not baseline-resolved (Figure 5-1(a)). Using cold-ion spectroscopy as a selective detector for a specific conformation, we observe that BK^{3+} forms at least three distinct conformational families in the gas phase, which were denoted I, II and III.[350] The distribution labeled I in Figure 5-1(a) corresponds to the most intense band in the UV spectrum of conformational family I. Conformational family II consists of two conformers, α and β. Their CV distributions are virtually identical and only that of conformer β is shown. Conformational families I and II are observed directly after ESI, while conformational family III requires some degree of collisional activation or sufficient time to be formed.

Normal bradykinin includes prolines in positions 2, 3 and 7, and the barriers for their *cis-trans* isomerization together with corresponding H-bonding re-arrangement are considered to be the origin of distinct conformational families.[70,173] Non-substituted bradykinin can potentially form $2^3=8$ different configurations, while if two out of three prolines are substituted by alanines, only two *cis-trans* isomers remain possible. Based on IMS data the proline in position 2 plays a special role in forming the peptide structure in the gas phase. Clemmer and co-workers examined a library of 58 model peptides with proline in the second position and found that 80% of them form two distinct conformational families, presumably because the barrier for *cis-trans* isomerization is lowered due to steric factors.[174] In the following section we explore the conformational space of triply protonated Pro 3&7 → Ala mutant of BK with a proline in position 2.
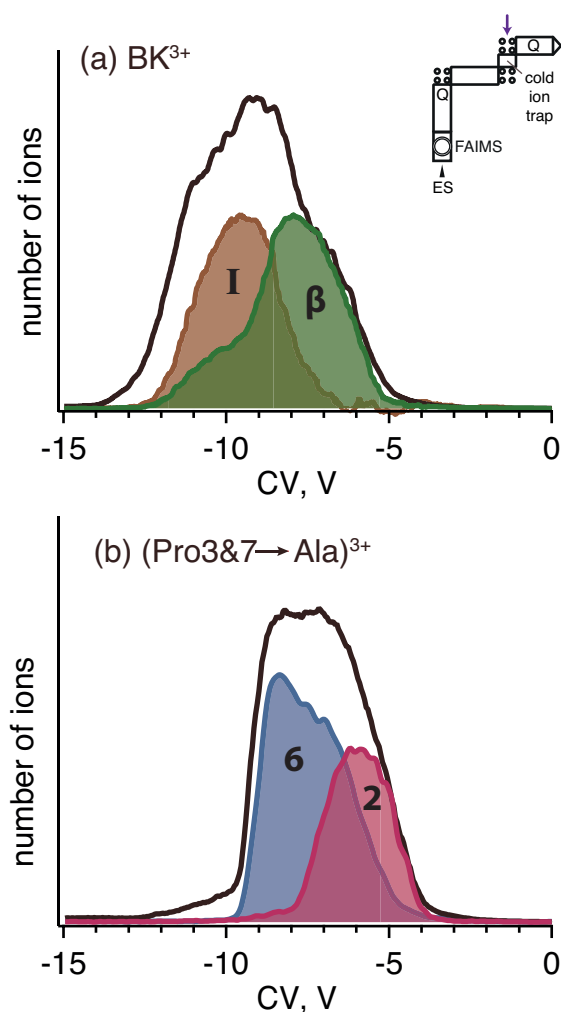
Figure 5-1. (a) CV distribution of BK^{3+} detecting all transmitted ions (black) and recorded *via* UV transitions unique for conformational families I (brown) and II (conformer β, in green). (b) CV distribution of Pro 3&7 → Ala mutant detecting all transmitted ions (black) and recorded *via* UV transitions unique for conformations 2 (pink) and 6 (blue).

5.3   Proline in positions 3 and 7 substituted by alanine

The BK mutant with prolines 3 and 7 substituted by alanine (Pro 3&7 → Ala) forms two distinct conformational families in the gas phase, as can be seen from its CV distribution (Figure 5-1 (b)). We observe two close maxima at -8.2 V and -7.1 V. A part of the distribution, which lies below ~-7 V, overlaps with that of non-substituted bradykinin, but a significant number of mutant ions appear at CV values between -7 V to -4 V.

In order to distinguish the conformational families of Pro 3&7 → Ala we set the CV on the shoulders of the distribution (-8.3 V and -6.0 V) and record the UV

photofragmentation spectra (Figure 5-2). Using the IR-UV double-resonance technique, we further obtain vibrational signatures for all intense bands in the UV photofragmentation spectra. We identify 3 main conformers in the spectrum at CV = -8.3 V (4, 5 and 6) and at least 5 conformers at CV = -6.0 V (1-3, 5 and 6). We separate the total CV distribution of the Pro 3&7 → Ala mutant into two main components by detecting selectively the photofragmentation signal of one conformation at a time, as shown in Figure 5-1(b). The distributions that correspond to structures 2 and 6 are presented; all other conformers fall into one of these two conformational families. Conformers 5 and 6 are present in both electronic spectra at CV=-8.3 V and at CV = -6.0 V, and the shape of their CV distributions suggests that they are formed from other structures upon collisional activation inside the FAIMS device and/or further in the instrument.
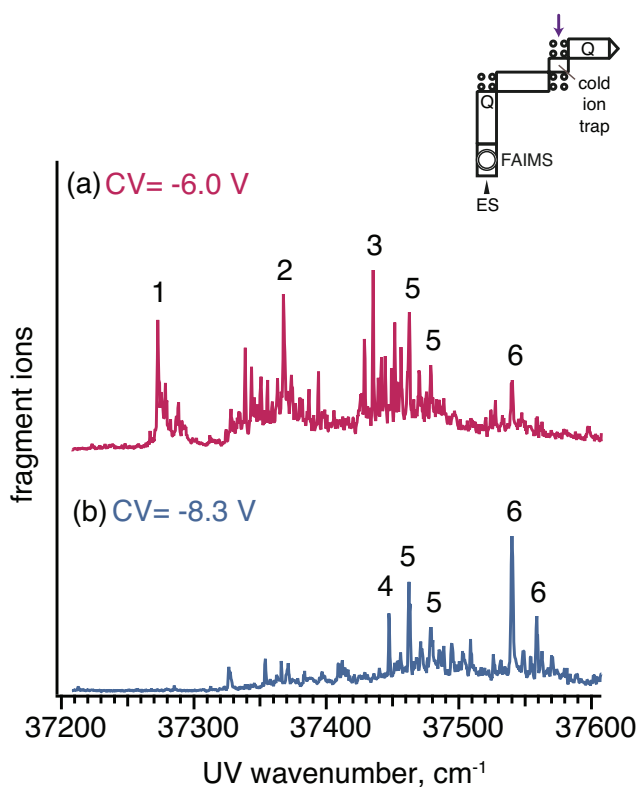
Figure 5-2. UV photofragmentation spectra of Pro 3&7 → Ala mutant obtained by setting the CV value to (a) -6.0 V; (b) -8.3 V. The conformers identified by IR-UV double-resonance spectroscopy are labeled with numbers.

127

The CV distributions of the Pro 3&7 → Ala mutant and non-substituted BK partially overlap. The maximum for conformational family II of BK lies at CV = -8.0 V, while for the mutant we find a conformational family with a maximum at CV = -8.3 V, which is close to -8.0 V, given the width of the peaks. We further compare the UV photofragmentation spectra taken at -7.0 V for BK and at -8.3 V for the mutant (Figure 5-3, this CV value for BK is used to avoid spectral contamination from family I). The spectra appear very similar: two major bands of conformer 5 correspond to those of conformer α, the same holds for conformers 6 and β. The UV spectrum of the mutant is blue-shifted by only 3.2 cm^{-1} compared to non-substituted BK, and most of the minor spectroscopic features are reproduced. However, there is an intense band in the photofragmentation spectrum of the mutant at 37439.2 cm^{-1}, which corresponds to conformer 4 and is not present in the case of non-substituted BK.
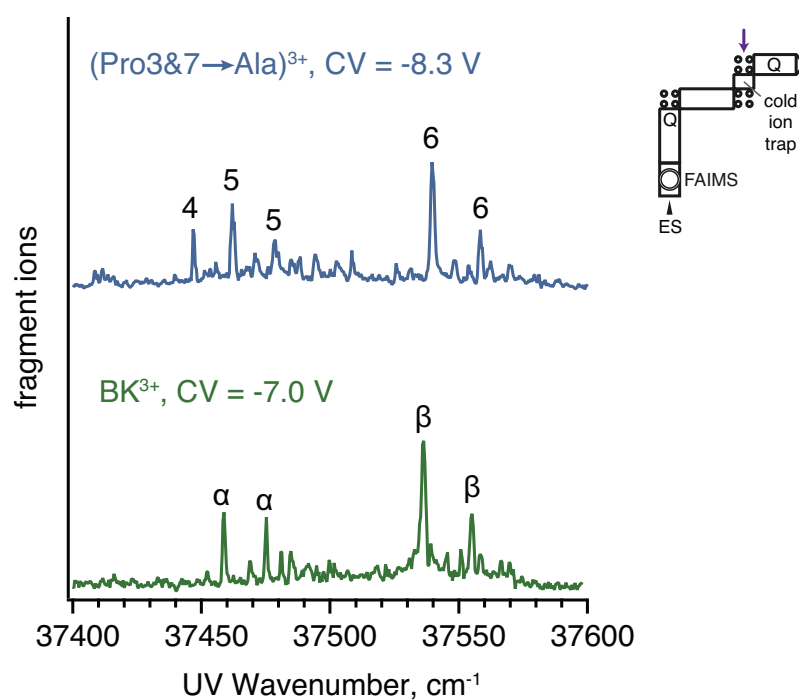
Figure 5-3. Comparison between UV photofragmentation spectra of Pro 3&7 → Ala mutant at CV = -8.3 V and conformational family II of BK^{3+} at CV = -7.0 V. The conformers identified by IR-UV double resonance spectroscopy are labeled.

The position of the UV bands and vibronic progressions suggests that both conformers α and β of non-substituted BK are present among molecular geometries adopted by Pro 3&7 → Ala. At the same time, the mutant also forms conformer 4, which

appears at the same CV as conformational family II of BK^{3+}, but has a different structure. To confirm these conclusions we compare the vibrational spectra of all conformers mentioned above, since the IR spectra reflect the hydrogen-bonding pattern and thus provide detailed information about molecular structure (Figure 5-4). Vibrational spectra of conformers 6 and β appear to be very similar, all bands are at the same position, except for the one at 3314 cm^{-1}, which is shifted by 7 cm^{-1}. It reflects a subtle change in the peptide structure, presumably the position of one of the side chains.[91,213] Replacing ^{14}N by ^{15}N in the backbone shows that this band corresponds to the NH stretch of serine (see section 4.5.3). Conformers 5 and α also have essentially the same geometry with a minor difference in the position of one of the side chains, most probably, due to the size difference between proline and alanine.
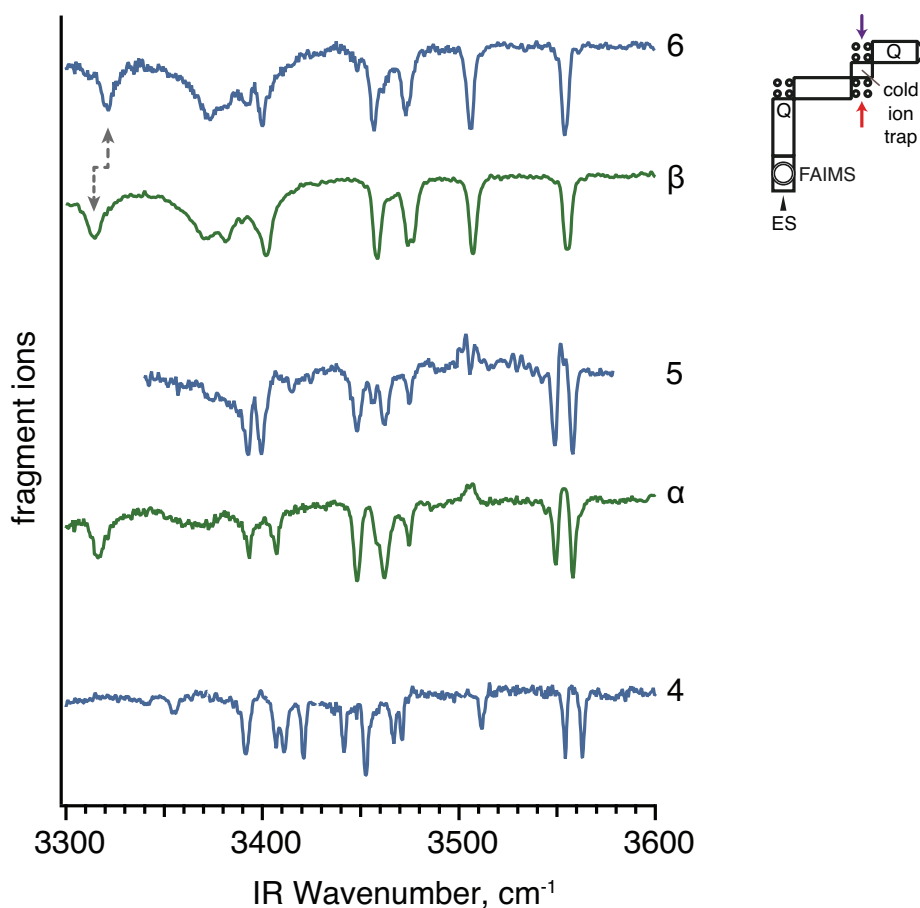
Figure 5-4. Comparison between vibrational spectra of conformers 4, 5 and 6 of Pro 3&7 → Ala mutant and conformers α and β of non-substituted BK^{3+}. Dashed line indicates the band of conformer 6 that shifts upon the mutation.

The structure of conformer 4, on the other hand, is substantially different from that of α and β. Its vibrational spectrum features two bands in the region higher than 3550 cm^{-1} that do not overlap with those of conformer α. There are at least 8 well-resolved sharp bands in the free-NH stretch region 3380-3490 cm^{-1} and only two weak bands below that, which suggests that there are fewer hydrogen bonds in conformer 4 than in conformers α and β. Indeed, upon collisional activation of the mutant ions, the relative intensity of conformer 4 goes down, proving that its free energy is higher than that of 5 and 6 (Figure 5-5).

To summarize, we observe that the Pro 3&7 → Ala mutant forms two conformational families in the gas phase. One of them overlaps with conformational family II of non-substituted BK, and conformers α and β of this family are reproduced by the mutant. It means they have prolines 3 and 7 in the *trans* position. The conformers 1-4 of Pro 3&7 → Ala are not accessible for non-substituted BK: structures 1-3 appear at the CV values higher than those recorded for BK, while conformer 4 has a distinct infrared signature, which reflects substantial differences in the structure.
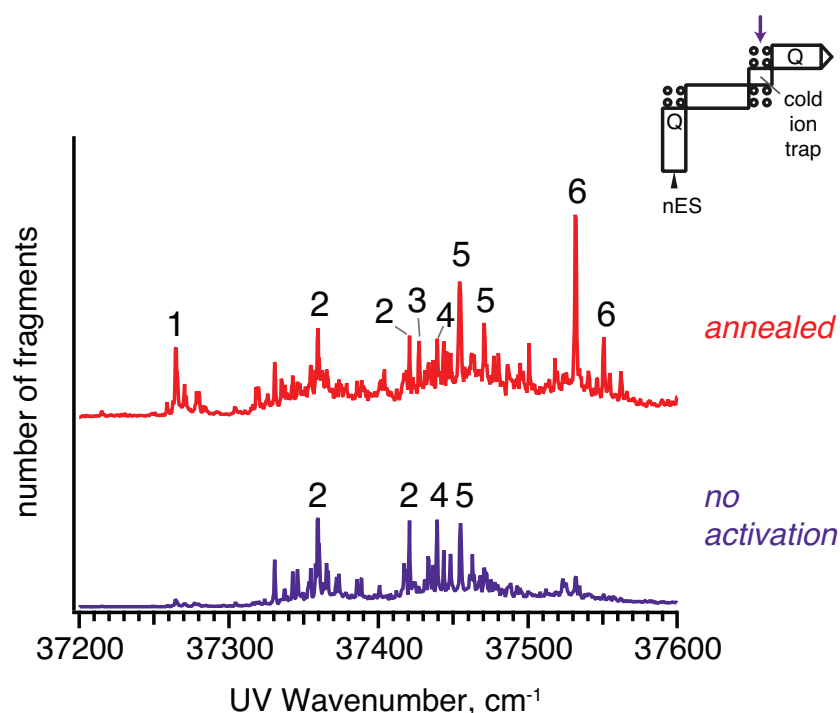
Figure 5-5. UV photofragmentation spectra of Pro 3&7 → Ala mutant with lowest collisional activation achievable (purple) and with high degree of collisional activation (red). The degree of activation is controlled by the amplitude of the RF voltage applied to the ion funnel electrodes.

We compare these findings with the information obtained by IMS. As reported previously by Clemmer and co-workers,[173] the drift-time distribution of BK has three resolved peaks, A, B and C, while the drift-time distribution of Pro 3&7 → Ala appears as one peak (Figure 5-6). We report here the CCS distribution of the mutant without adjusting for the size difference between proline and alanine, which would shift the distribution by 4.9 Å^2 towards larger CCSs. If this factor were taken into account, the peak of the mutant would fully overlap with conformational family B of non-substituted BK. However, given that the structures of conformers 5 and 6 are virtually the same as that of α and β, adjusting for the size of the average size of amino acid is not necessarily justified.

The collisional cross-section distribution of the Pro 3&7 → Ala mutant partially overlaps with family B of non-substituted BK (see Figure 5-6). The structure of conformers α and β of non-substituted BK is the same as of conformers 5 and 6 of the mutant respectively, which means that their cross-sections are close. Thus these conformers represent the overlap between the two CCS distributions depicted in Figure 5-6. We conclude that conformational family II separated by FAIMS corresponds to family B isolated by DT-IMS. It is confirmed by similar behavior of these families upon collisional activation.[350]
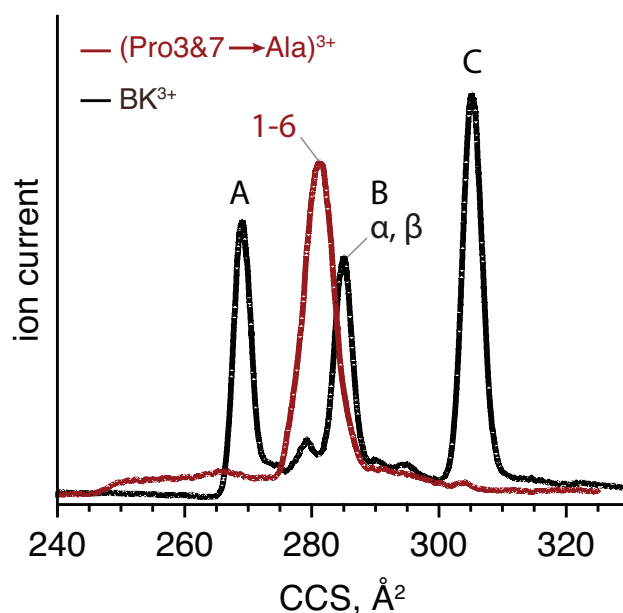
Figure 5-6. Drift time distribution of Pro 3&7 → Ala mutant compared to non-substituted BK^{3+}.

Importantly, we observe that underneath one peak in the Pro 3&7 → Ala IM distribution there are more conformations than within family B/II of non-substituted BK. An additional conformational family appears upon mutation, which might correspond to the other isomer of the peptide bond preceding proline in position 2 or to a significantly different hydrogen-bonding pattern.

5.4 Proline 2 substituted by alanine

We previously suggested that family III of BK[3+] separated by FAIMS corresponds to family C isolated by DT IMS based on their similar behavior upon collisional activation: both consist of stable structures that are produced mainly in the gas-phase.[350] Results obtained by Clemmer and co-workers, on the other hand, show that the CCS of family C of BK[3+] is close to that of Pro 2 → Ala mutant.[173] The drift-time distributions of both ions are shown in Figure 5-7.
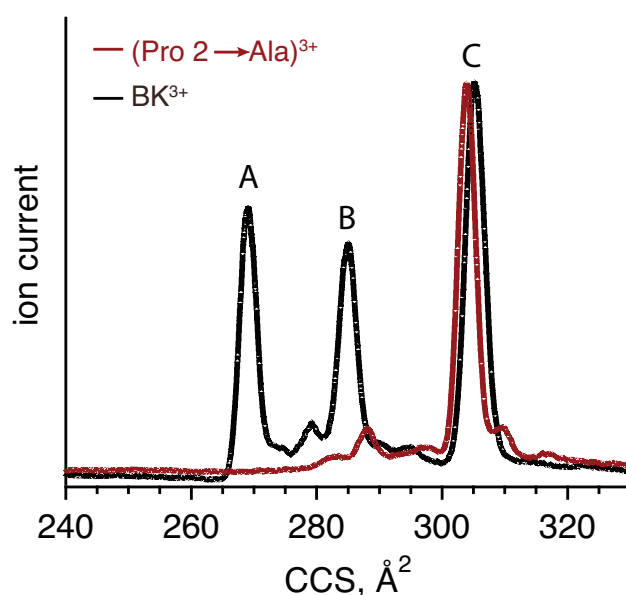
Figure 5-7. Drift time distribution of Pro 2 → Ala mutant compared to non-substituted BK[3+].

We further compare the UV photofragmentation spectra of the Pro 2 → Ala mutant with that of family III of BK[3+] in order to support the general conclusions made for Pro 3&7 → Ala. We also show how this mutant can be used to simplify the UV spectroscopic signature and to obtain conformer-selective vibrational spectra.

Conformational family III of BK^{3+} cannot be fully isolated by FAIMS in the same way as families I and II are. This is explained by the origin of this family: we observe that it is produced mostly in the gas phase as a result of collisional activation. Its UV photofragmentation spectrum is thus obtained as a difference between the UV spectrum of BK^{3+} recorded at CV = -9V and two other conformational families (see section 4.4). The result of subtraction is shown in Figure 5-8. We observe a similarity between the UV photofragmentation spectrum of Pro 2 → Ala and the spectrum of family III of non-substituted BK. Sharp peaks, labeled 1, 1', 2, and a broad feature, which presumably consists of numerous close bands, overlap in the two electronic spectra. At the same time the spectrum of the Pro 2 → Ala contains many more lines at 37460-37500 cm^{-1}. It means that qualitatively the result of substituting proline in position 2 by alanine is similar to that of substituting prolines 3 and 7: a part of conformers of native BK are reproduced, but at the same time a number of new ones are formed.
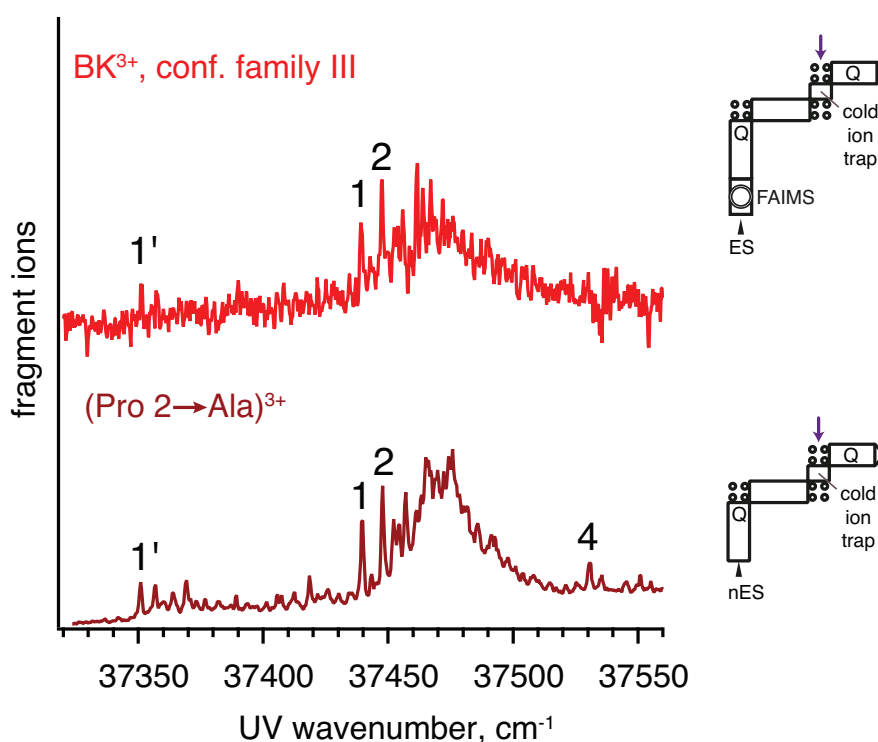
Figure 5-8. Comparison between the UV photofragmentation spectra of conformational family III of BK^{3+} and the Pro 2 → Ala mutant. The conformers of interest are labeled: 1', 1 and 2 have overlapping UV transitions and likely have similar structure.

In the region 37430 cm^{-1} to 37600 cm^{-1} the UV photofragmentation spectrum of BK^{3+} is dominated by conformers α and β, although upon collisional activation they partially interconvert to conformational family III (see section 4.4). Thus recording the vibrational spectra of family III in the case of non-substituted BK is challenging: when we fix the UV laser on a band of family III and shine the IR laser, which is in resonance with a vibrational transition of conformer α or β, UV absorption lines of the major conformer broaden and the gain signal produced by this broadening is comparable with the depletion of the band on which the UV laser is set. Proline to alanine substitution allows us to overcome this obstacle, since the electronic spectrum of Pro 2 → Ala corresponds only to conformational family III and is not "contaminated" by other families. We thus recorded the IR spectroscopic signatures of conformers 1, 1' and 2 (Figure 5-9), and we assume that their structure is similar to the corresponding conformers of non-substituted BK. It is remarkable that even these geometries with high CCS do not have a free OH of the serine side chain, which would produce absorption signal above 3600 cm^{-1}. Spectra of conformers 1 and 1' differ only by the position of one line, while the spectrum of conformer 2 has several unique bands, but also share spectroscopic features at 3337 cm^{-1}, 3390 cm^{-1}, 3450-3465 cm^{-1}, and 3547-3564 cm^{-1} with conformer 1. This suggests that they indeed form a conformational family defined as a set of structurally similar conformers.
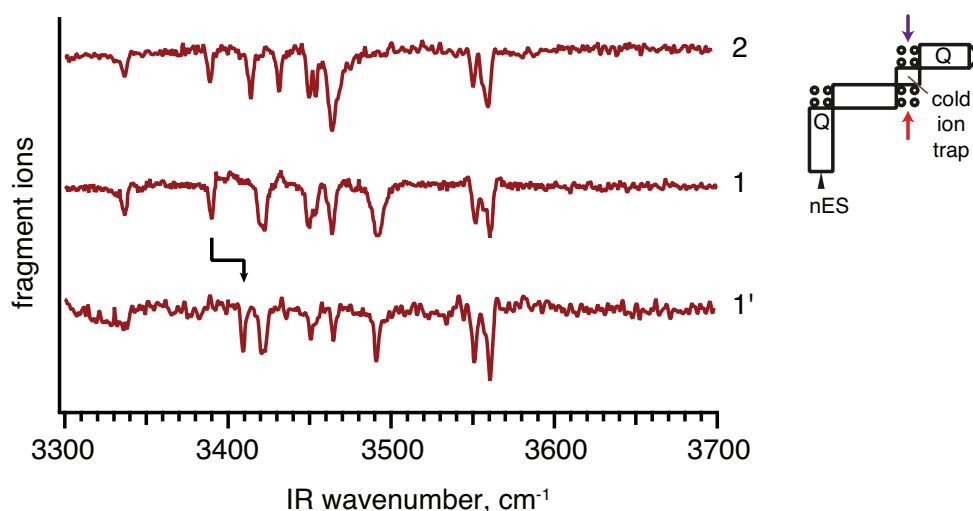
Figure 5-9. Vibrational spectra of Pro 2 → Ala mutant on the UV transitions that overlap with those of non-substituted BK^{3+}. Spectra of conformers 1 and 1' differ by position of one band, shown with an arrow.

5.5   Why do we observe additional conformers?

The local structure of a peptide is determined by three dihedral angles, usually denoted as φ, ψ, and ω (Figure 5-10). Up to now we have discussed the angle ω, which can take two discrete values, around 180^0 in the *trans* conformation or around 0^0 in the *cis*, due to partial double-bond character of the peptide bond. Obviously, the conformational preferences of proline and alanine relative to other dihedral angles are also important.

Energetically allowed dihedral angles are usually depicted as a Ramachandran plot, introduced by Ramachandran and co-workers in 1963.[378] Since then, the number of high-resolution protein structures has been growing rapidly, and Ramachandran plots for every amino acid in different conditions have been generated.[160,379-381] They are mainly used for validation of protein structure obtained by X-ray crystallography.[382-384]
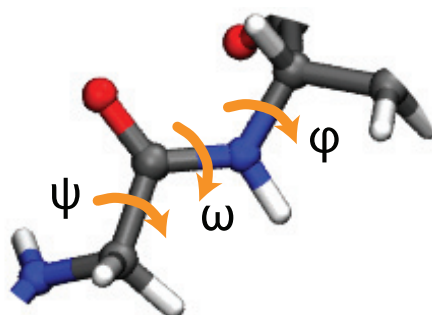
Figure 5-10. The peptide bond and the backbone dihedral angles.

We would like to compare the conformational preferences of proline and alanine. Chakrabarti and co-workers analyzed a database of 408 polypeptide chain structures, which included 7050 alanine and 3966 proline residues,[381] later Lovell and co-workers performed a similar study of a 500-structure database.[160] They defined torsion angles for each amino acid and presented all data points in Ramachadran plots (Figure 5-11 and Figure 5-12(c)).
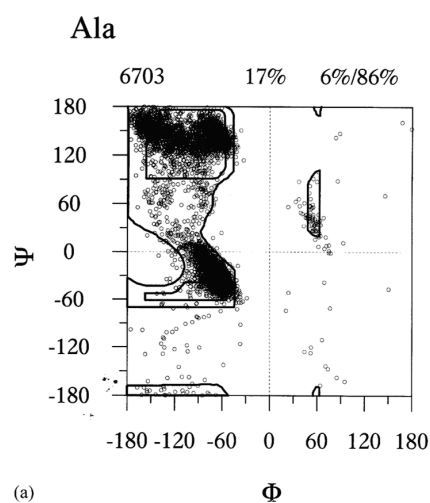
Figure 5-11. The φ, ψ distributions of Ala residue superimposed on the respective Ramachandran plot. Reproduced with permission.[381]
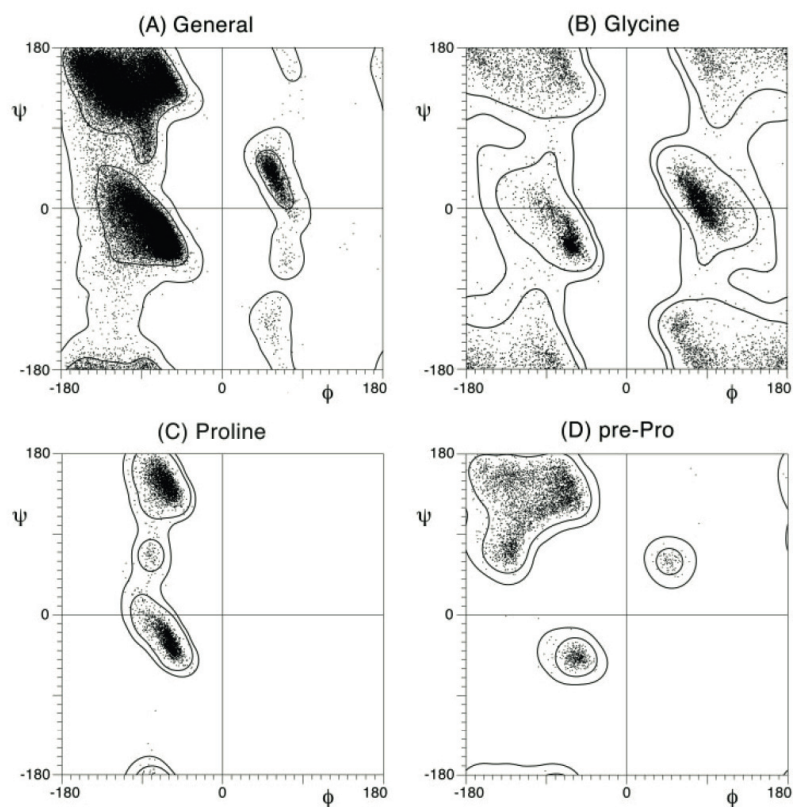
Figure 5-12. φ, ψ angle distributions for 97,368 residues from the 500-structure high-resolution database, along with validation contours for favored and allowed regions. (a) The general case of 81,234 non-Gly, non-Pro, non-prePro residues. (b) The 7705 Gly residues, shown with twofold symmetrized contours. (c) The 4415 Pro residues with contours. (d) The 4014 pre-Pro residues (excluding those that are Gly or Pro) with contours. Reproduced with permission.[160]

136

Clearly, the alanine backbone can occupy a significantly larger part of the conformational space and adopt structures that are not achievable for the proline residue due to the nature of its side chain. Specifically, the proline backbone is fixed around $\varphi \sim -60^0$, while alanine can be found with $\varphi < -120^0$ and $\varphi \sim +60^0$. This flexibility allows a peptide in which proline is substituted by alanine to adopt conformations that could not be achieved by the original peptide. In general, these new conformations may have low relative energy, and their CCS might be not related in any way to the IM distribution of the original peptide.

A residue preceding proline is also conformationally restricted compared to an average amino acid (Figure 5-12 (A) and (D)), which means that proline rigidity influences the surrounding amino acids. When proline is substituted by alanine, much larger conformational space becomes available. In the case of BK, we observe nearly identical conformations for the substituted and non-substituted peptide, but in general this does not have to be the case. It is probable that the conformations of the alanine-containing peptide with the torsion angles φ, ψ forbidden for proline are lower in energy than the original structures of the proline-containing molecule.

Analyzing the differences between (φ, ψ)-distributions of natural amino acids, Chakrabarti[381] proposed a matrix of conformational similarity based on the backbone torsion angles. The similarity index ranges between 0 and 1, and for the proline–alanine pair equals 0.44, which means that their conformational preferences are significantly different. Anderson and co-workers[385] took this approach one step further by creating a table of substitutional priority. It shows that amino acids highly similar to proline in terms of conformations are the ones with bulky side chains, such as isoleucine, valine, methionine, tryptophan and leucine. If all amino acids are ranked according to their similarity to proline, alanine is 12th out of 19. Glycine is the least similar to proline natural amino acid, which can be illustrated by comparing Figure 5-12 (B) and (C). Such a substitutional priority table can help us to choose a conformationally closer substitute to proline than alanine, e.g. isoleucine or valine, in order to have a higher chance of preserving the original geometry of the peptide under investigation and to avoid generating new conformations.

5.6 Conclusions

After a series of studies performed by Clemmer's group,[165,173,174] it has been often assumed that multiple peaks in the IM distribution of peptides mainly result from *cis-trans* isomerization of proline and that the identity of each peak can be accessed using point mutations.[142,169,176,386,387] Recently Lietz and co-workers raised a question "Multiple gas-phase conformations of proline-containing peptides: Is it always *cis/trans* isomerization?"[386] In most cases IM peaks were attributed to *cis-trans* conformers of proline using additional measurements and such assignment was sufficiently justified. However, the general answer to the question cited above is, most probably, "no".[386]

In the title of this chapter we pose a related question: "Can proline mutations be used to assist structure determination?" Considering the experimental evidence, the answer is "yes, but more than a single CCS is needed to support the assignment". We observe that a part of the initial conformational distribution is reproduced both by Pro 3&7 → Ala and Pro 2 → Ala mutants, as expected. Not only the CCS of each mutant coincides with one peak in the original IM distribution, but also the IR spectroscopic signatures of certain conformers of the mutants are almost identical to those of non-substituted BK. This indeed allowed us to establish that conformers α and β of native BK have prolines 3 and 7 in the *trans* conformation and that in the lowest-energy gas-phase conformations proline 2 is *trans*. However, we also showed that due to the high flexibility of the alanine backbone compared to proline, the mutants adopt additional conformations, which were not achievable for the non-substituted BK. The nature of these conformers is, generally, sequence-dependent and can be misleading for the assignment of the initial IM distribution. Figure 5-13 shows a schematic representation of the conformational space of BK and its mutant. The conformational distributions overlap, but neither one is part of the other.
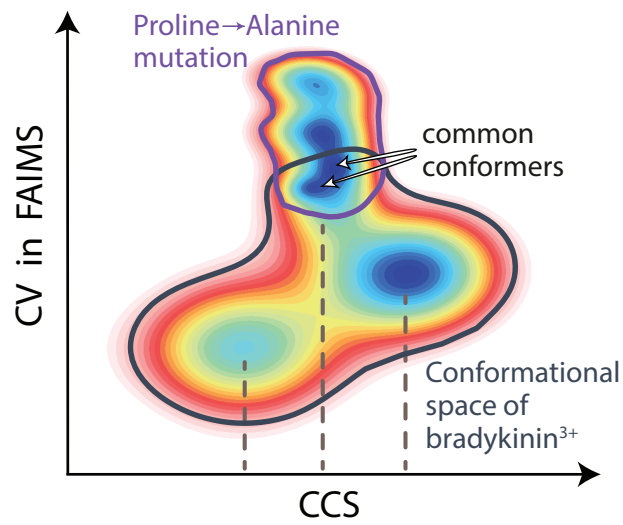
Figure 5-13. Schematic representation of the conformational space of triply protonated BK and of the Pro 3&7 → Ala mutant. The color corresponds to the free energy.

Chapter 6.
Conclusions and outlook

Functions of proteins and peptides are closely related to their structure. While enzymes and fibrous, transport and receptor proteins mostly have a well-defined, rigid 3D structure, a large number of biological molecules are partially or fully intrinsically disordered, which makes them a challenging target for the traditional techniques of structure determination such as NMR and X-ray crystallography. Gas-phase techniques, on the other hand, can be mass- and conformer-selective, which simplifies the structure analysis of conformationally heterogeneous systems. There are a number of issues that prevent gas-phase methods from finding broader use, however. One of them is the question of to which extent the structure of a biological molecule is preserved when transferring it from solution to the gas phase *via* electrospray. In order to answer this question fully, one needs to identify which structural elements are preserved in the electrospray process and which ones are scrambled.

It is well established that the electrospray process often generates a non-equilibrium conformer distribution, which contains structures that are kinetically trapped behind high barriers on the free energy surface. These structures are of particular importance, as they provide a link between the solution state of the molecule and its lowest-energy, gas-phase conformations. In this thesis we have focused on the identification and characterization of such conformations.

Cis/trans isomerization of the peptide bonds has one of the highest barriers in the molecule, and the corresponding states have a high chance to be preserved in the electrospray process. With that in mind, we studied several model systems with large number of proline residues, for which *cis* conformers are comparable in energy to *trans,* and, as a consequence, an appreciable percentage of *cis* conformers is typically

observed. We see this as the first step in investigating how different structural elements are preserved or distorted in the electrospray process.

The first example presented in this thesis, the bradykinin 1-5 fragment, includes two proline residues, which adopt mostly *trans-trans* conformation in solution. Curiously, in the gas phase we observe three distinct conformational families. The extended structures are kinetically trapped and preserve the solution-phase conformation of the peptide bonds, i.e. all-*trans*. Two other conformational families are formed in the gas phase and/or during the electrospray process through isomerization of one of the prolyl-peptide bonds to give *cis-trans* and *trans-cis* configurations. Although according to quantum chemical calculations the *cis-cis* conformation is as low in energy as the *cis-trans* and *trans-cis*, we do not observe it in our experiments. To summarize, the doubly protonated bradykinin 1-5 fragment has been characterized in the gas phase and in solution, and the results confirm that proline *cis/trans* isomerization state can be preserved in the electrospray process. This result is somewhat surprising for a molecule of this size, and we attribute it to the height of the *cis/trans* isomerization barrier. Other structural elements might require a greater number of intermolecular interactions to be preserved in the electrospray process.

The case of bradykinin is more complex than that of its 1-5 fragment. First, it has four additional bulky amino acids, which increases the time required for quantum chemical modeling. Second, it contains three prolines, leading to 8 possible conformational classes instead of 4. Here we focused on the +3 charge state and extensively characterized the low-energy and the kinetically trapped structures spectroscopically. In agreement with the previously published ion mobility data, we identified three conformational families, one of which is kinetically trapped, since it completely disappears upon annealing. The relative abundances of two other conformational families also change upon collisional activation. Moreover, using double-resonance spectroscopy, we identify within each conformational family a number of conformations with different hydrogen-bonding patterns.

While the kinetically trapped conformations are important for the insight they provide into the solution-phase structure of the molecule, the gas-phase conformational distribution can shed light onto the state of the molecule in non-polar media such as a cell membrane. Further research is needed to confirm experimentally the hypothesis that measurements in vacuum can be used to predict the behavior of a molecule in

142

biologically relevant media with low electrical permittivity. We could potentially characterize the transition from the aqueous environment to a lipid membrane for a model antimicrobial peptide and compare it to what can be determined in vacuum.

We usually attribute kinetic trapping to the result of evaporative cooling and the short lifetime of the nanodroplets in the electrospray process. However, the distribution that we observe in the ion trap reflects an entire sequence of changes in the conditions that the ions are exposed to. First, kinetic trapping by definition means that there is a slow interconversion of the ions from the high-energy structures to the low-energy ones. This reaction can be accelerated intentionally by collisional activation of the ions in the ion funnel or in the activation region of the drift tube. The rate of interconversion can be directly measured by monitoring the drift-time distribution as a function of the pre-trapping time. The molecules might also be unintentionally activated in regions where the pressure is low enough to give ions time to accelerate in the electric field and high enough for at least one collision to occur. As we have shown for the *trans-trans* conformational family of BK[1-5], at room temperature the low-energy conformations are more extended than at the temperature of the trap ($\sim$10 K). Knowing that, we would expect certain structural re-arrangement during bath gas cooling in the cryogenic trap. Nevertheless, we recorded the spectroscopic signature of the extended conformations and not of the compact ones, which suggests that cooling in collisions with cold helium is sufficiently fast for kinetic trapping to happen in the cryogenic trap. Additional experiments have to be performed to determine the effective temperature that corresponds to the conformational distribution in the trap.

In its native environment, a biological molecule is in a dynamic protic equilibrium with the solvent and may not have a well-defined charge state. In the gas phase, on the other hand, every peptide or protein appears in a number of distinct charge states, and if we were to compare a conformational distribution in solution to that in the gas phase, all charge states have to be taken into account. For BK, for instance, it has been shown that kinetically trapped structures exist both for the +2 and +3 charge states. What is the relationship between them? Do the conformational distributions of different charge states uniformly sample the original conformational ensemble in solution or does a certain structural family populate a preferred charge state? (The latter hypothesis is supported by the observation that the disordered unfolded proteins usually populate higher charge states than the compact folded ones.) These are examples of questions

that still have to be answered in order to achieve a better understanding of the relationship between gas-phase and solution-phase conformational preferences of biological molecules.

From the experimental standpoint, we have combined linear and field-asymmetric IMS with cryogenic spectroscopy. The results presented in this thesis show that details of the molecular structure cannot be resolved using IMS alone, since it provides only the information about the overall shape of the molecule. If we are interested in the backbone conformation of a peptide, we have to employ complementary techniques, such as vibrational spectroscopy. This was clearly illustrated by our investigation of the proline-to-alanine mutants. In general, such substitution allows one to identify the conformation of the prolyl-peptide bonds in certain molecular geometries, as suggested by the IMS data. However, we also observed a number of additional conformations, which are induced by the high flexibility of the alanine backbone compared to the proline. These new conformations can potentially scramble the assignments, if CCS distribution alone is used for the analysis. Vibrational spectroscopy, on the other hand, produces a spectroscopic signature unique for a particular hydrogen-bonding pattern, which can be used for a detailed comparison of the conformational space of the mutant and the native peptide. We also suggest using the table of substitutional priority to pick amino acids that tend to adopt similar geometries to proline.

Another issue that holds back the development of gas-phase vibrational spectroscopy as a tool for structure determination is a need for high-level quantum chemical simulations to interpret and complete the experimental results. Currently computer modeling has limited capabilities for treating larger biological systems. In this regard, the measurements of the vibrational frequencies in vacuum are particularly valuable, as they allow direct comparison with the calculations. In this study we showed how CCSs and IR spectra are used to guide and verify first-principles simulations, giving rise to a new type of symbiosis between theory and experiment. We emphasize that their integration is important for the development of new specialized computational methods, which would allow us to investigate larger peptides and proteins. Already in the size range of 20-50 amino acids there are a number of biologically important compounds (e.g. amyloids) that present a challenge to solution-phase techniques. The detailed study of the smaller model systems presented here

brings us one step closer to understanding these complex molecules in the gas phase, and, through simulation of the kinetically trapped structures, in solution.

In general, we followed a bottom-up approach: we chose two relatively small model systems and combined advanced experimental and theoretical methods to characterize the conformational preferences of the prolyl-peptide bonds and to obtain additional structural information. We believe that a number of detailed studies of peptides with different sequences and molecular geometries (beta-sheets, different types of turns, abnormally high *cis* percentage, etc.) are needed to create enough material for a complete answer to the question "Which structural elements are preserved in the electrospray process and to what extent?" This work is one of the first steps towards a database of atomic-resolution peptide structures in the gas phase.

Acknowledgements

This work would not have been possible without the Laboratory of Molecular Physical Chemistry and especially its Head and the supervisor of this thesis, Prof. Thomas Rizzo. I would like to thank Tom for accepting me to his subgroup and giving me all independence and resources a student can hope for. Thanks to his constant support and guidance I developed my own scientific intuition, collaborated with other groups and presented the results at many exciting conferences. Tom's approach to science and his desire to help students develop both academically and personally have influenced a lot my own value system.

This work has been defended in front of the experts, whom I would like to thank for accepting the invitation, reading this manuscript, asking challenging questions at the defense and for the discussions we had at the conferences. I also would like to thank them for the informal interactions and sharing their personal experience with the "young generation".

When I joined LCPM, I did not have any experience in working with home-built instruments. Aleksandra was the one who introduced me to the machine and the way the group functions, and she became my close friend since then. I am grateful for all the moments we have spent together! In the first years of my PhD I received a lot of help and support from Nils in battling the lasers, from Yiorgos in urgent FAIMS installation, from Oleg in never ending improvement of the laser alignment, from Xiohang in adjusting to the group dynamics, from Antoine in computer issues… A special "thank you" goes to Marianne, the "Swiss Mum" of the LCPM kids. I appreciate a lot the help, and even more the friendship she offered me. I cannot overemphasize the importance of the discussions we had with Mike throughout the time we were working together. We exchanged ideas and shared the successes, and Mike has helped me to go through all difficulties. We have worked a lot together with Chiara, and her decisive character and endless optimism have made the days in the lab both more fun and

productive. I also would like to thank Ulrich for being very encouraging while I was writing this manuscript and for his exceptional sense of humor.

In the last couple of years we had a subgroup meeting almost every week, and this is one more way how the bio-group has contributed to the content of this thesis. Over pistachios and cookies we went through all ideas, data, presentations and misconceptions. I was happy to be part of such group and I am grateful for all the support and criticism I got!

A substantial part of this thesis is the "computer experiment", which was done under supervision of Dr. Carsten Baldauf. Our collaboration added a new dimension to this work, not only because of the scientific value of the results, but also thanks to Carsten's motivation and constructive approach. I appreciated the opportunity to discuss the results with him ASAP and I would like to thank Carsten for taking so much time for our conversations and for accepting my doubts and certain perfectionism. I am also very grateful for the possibility to visit FHI in Berlin, to work together with Franzi, Markus, Mateusz and others. I spent a great time there, both professionally and personally, and I have fall in love with Berlin!

All people whom I have mentioned above (and many more "surface", "helium" and "bio" lab members) are were only my colleagues, but also close friends, without whom I wouldn't have had that fulfilling PhD life. I would like to thank all the past and present members of LCPM for the time we have spent together. I think it's a great privilege to work in such atmosphere and I can only wish that it remains this way for the next generation of LCPM members. We shared so much fun time: hiking, biking, skiing, bar-hopping, running races, wine tasting, rafting, swimming, dancing, BBQ-ing… And of course, we were having lunch together every day and even established new group traditions! Despite the controversy, I should be happy that I happened to sit next to the coffee machine: I could learn every day something new from the people passing by!

I got inspiration not only from the lab members, but also from my friends living and working far away. Whether they are in Moscow, Chelyabinsk, Paris, Canada, Marseille, Berlin, Zurich, Macedonia… or (lucky coincidence!) Lausanne, I can always count on them and feel their presence in my life. Having friends and acquaintances from all over the world opened for me new ways of looking at ordinary things.

Research work implies many disappointments on the way to a worthy result. Luckily, in such moments other projects kept me motivated, and therefore I

Curriculum Vitae

EDUCATION

PhD candidate,
Laboratory of Physical Molecular Chemistry   2011 – 09.2016
Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland.
Under supervision of Prof. T.R. Rizzo.

- Award for the **best poster presentation**,   07.2013
  Gordon Research Conference on Biological Ions in the
  Gas phase & in Solution, Holderness NH, USA.

MSc with Honors (4.8/5.0)   2009 – 2011
Department of Physics of Supramolecular Systems and Nanophotonics,
Moscow Institute of Physics and Technology State University (MIPT), Moscow, Russia.

- "Sensor for volatile hydrocarbons" – a project   2009 – 2011
  funded by an award of 10.000 euros by MIPT Innovation Center.
- Project Management Expert",   06.2010
  a certification based on PMI PMBOK.
- 1st prize at the 4th Internet-Olympiad in Nanotechnology,   03.2010
  creative competition, Moscow, Russia.

BSc with Honors (4.95/5.0)   2005 – 2009
Applied Physics and Mathematics, Faculty of Molecular and Biology Physics, MIPT.

- **Scholarship for academic excellence** awarded   2007 – 2008
  by the Foundation for the Development of Innovation Education, Russia.

3rd Prize at the All-Russian Physics Olympiad   2004, 2005
for high school students, Russia

RESEARCH METHODS AND SKILLS

- Laser spectroscopy (UV, IR), optics
- Mass-spectrometry
- Ion mobility spectrometry
- Ultra-high vacuum technology
- Basics of Python
- Quantum chemical calculations
- Absorption and fluorescence spectroscopy

LANGUAGES

Russian – native, English – C1, French – B2, Serbian – B1/B2

RESEARCH EXPERIENCE

Laboratory of Physical Molecular Chemistry,   2011 – 2016
EPFL, Lausanne, Switzerland

- Cryogenic ion spectroscopy for structure determination of biological molecules in the gas phase.
- Using and developing two home-built tandem mass-spectrometers.
- Collaborations with experts in the NMR and a group specialized in quantum chemical calculations.

Laboratory of Structure and Dynamics of   2008 – 2011
Molecular-Organized Systems,
Photochemistry Center of Russian Academy of Science, Moscow, Russia.

- In the frame of an "artificial nose" project, established a workflow to prepare and test materials that absorb volatile aromatic compounds.
- Although the laboratory had limited experience in spectroscopy, performed the initial measurements of concentration of the odour using a commercial fluorescent spectrometer.

TEACHING EXPERIENCE

Certified Instructional Skills Workshop (ISW) facilitator,   2015 – now
Teaching Support Center at EPFL (CAPE).

- Providing a 3 days intensive training for the instructors with different levels of experience, from teaching assistants to professors.

Teaching assistant for Quantum Chemistry,   2012 – 2016
Spectroscopy and General Physics courses, EPFL.

          Special prize for exceptional teaching, Chemical Section, EPFL.   2015

Physics teacher, School of Physics and Technology, MIPT.   2006 – 2011

- Preparing a group of 20 high school students for the entrance exams to MIPT.
- Training high school students for Olympiads of different levels, including International.

ACTIVITIES

Representative of Switzerland in the International Organization   2012 – now
Committee of International Physicists' Tournament (IPT).

- A team leader of the Swiss team taking part in the IPT for 4 years.
- As a jury member, judged the performance of students in solving experimental problems.

Member of the Local Organizing Committee of IPT: 2012, 2013
an international event for 100 participants from 11 countries

- A member of a team of 6 students who organized the IPT for the first time at EPFL.

TALKS AT INTERNATIONAL CONFERENCES

- 64th ASMS Conference on Mass Spectrometry and Allied Topics, June 5-9, 2016, San Antonio, USA.
- 9th Conference on Isolated Biomolecules and Biomolecular Interactions, April 10-15, 2016, Oxford, UK.
- Strasbourg Workshop on Native Mass Spectrometry and Ion Mobility, April 8-9, 2014, Strasbourg, France.

POSTERS AT INTERNATIONAL CONFERENCES

- Fall meeting of the Swiss Chemical Society, September 4, 2015, Lausanne, Switzerland.
- Gordon Research Conference "Gaseous ions: Structures, Energetics and Reactions", February 22-27, 2015, Galveston, TX, USA.
- 8th Conference on Isolated Biomolecules and Biomolecular Interactions, May 18-23, 2014, Porquerolles, France.
- Fall meeting of the Swiss Chemical Society, September 6, 2013, Lausanne, Switzerland.
- Gordon Research Conference "Biological Molecules in the Gas Phase & in Solution", July 21-26, 2013, Holderness, NH, USA.

PUBLICATIONS

- Conformations of prolyl–peptide bonds in the bradykinin 1–5 fragment in solution and in the gas phase.
Voronina, L., Masson A., Kamrath, M., Schubert, F., Clemmer, D., Baldauf C. & Rizzo, T. R.
J. Am. Chem. Soc., **138**, 9224–9233, (2016).
- Spectroscopic studies of kinetically trapped conformations in the gas phase: the case of triply protonated bradykinin.
Voronina, L. & Rizzo, T. R.
Phys. Chem. Chem. Phys. **17**, 25828-25836, (2015)
- Transfer of electronic excitation energy from naphthalene to fluorophore indicator on the surface of silica microspheres with covalently or physically bound cyclodextrins.
Voronina, L. V., Meshkov, B. B., Tsybyshev, V. P., Alfimov, M. V. & Livshits, V. A.
Nanotechnologies in Russia **8**, 603-611, (2013).
- Complexation of a gaseous spin probe with cyclodextrins bound to the silica microspheres: Molecular dynamics of the complexed probes and the effect of aromatic hydrocarbon vapors.
Ionova, I. V., **Voronina, L. V.**, Meshkov, B. B., Alfimov, M. V. & Livshits, V. A.
Nanotechnologies in Russia **8**, 592-602, (2013).
- Studies of possibility of detecting aromatic hydrocarbons in gas phase on the basis of fluorescence of host-guest complexes with cyclodextrins immobilized on silica microspheres.
Voronina, L. V., Livshits, V. A. & Alfimov, M. V.
Nanotechnologies in Russia **6**, 444-455, (2011).
- Sections "Computational chemistry" and "Förster resonance energy transfer" in "The dictionary of nanotechnological and related terms"
Voronina, L. V.
PhysMathLit, Moscow, 2010.

PUBLICATIONS IN PREPARATION

- How does substituting proline by alanine change the conformational space of a peptide?

153

Bibliography

1   Voet, D. & Voet, J. G. *Biochemistry, 4th edition*. (Wiley, 2011).

2   Severini, C., Improta, G., Falconieri-Erspamer, G., Salvadori, S. & Erspramer, V. The tachykinin peptide family. *Pharmacol. Rev.* **54**, 285-322, (2002).

3   Catt, K. J., Harwood, J. P., Aguilera, G. & Dufau, M. L. Hormonal regulation of peptide receptors and target cell responses. *Nature* **280**, 109-116, (1979).

4   Axelrod, J. & Reisine, T. D. Stress hormones: Their interaction and regulation. *Science* **224**, 452-459, (1984).

5   Tatar, M., Bartke, A. & Antebi, A. The endocrine regulation of aging by insulin-like signals. *Science* **299**, 1346-1351, (2003).

6   Duquesne, S., Petit, V., Peduzzi, J. & Rebuffat, S. Structural and functional diversity of microcins, gene-encoded antibacterial peptides from enterobacteria. *J. Mol. Microbiol. Biotechnol.* **13**, 200-209, (2007).

7   Willey, J. M. & van der Donk, W. A. Lantibiotics: Peptides of diverse structure and function. *Annu. Rev. Microbiol.* **61**, 477-501, (2007).

8   Cotter, P. D., Hill, C. & Ross, R. P. Bacteriocins: Developing innate immunity for food. *Nat. Rev. Microbiol.* **3**, 777-788, (2005).

9   Nicolas, P. Multifunctional host defense peptides: Intracellular-targeting antimicrobial peptides. *FEBS J.* **276**, 6483-6496, (2009).

10   Hong, M. & Su, Y. Structure and dynamics of cationic membrane peptides and proteins: Insights from solid-state NMR. *Protein Sci.* **20**, 641-655, (2011).

11   Hall, J. M. Bradykinin receptors: Pharmacological properties and biological roles. *Pharmacol. Ther.* **56**, 131-190, (1992).

12   Dimitrov, D. S. Virus entry: Molecular mechanisms and biomedical applications. *Nat. Rev. Microbiol.* **2**, 109-122, (2004).

13   London, N., Raveh, B. & Schueler-Furman, O. Peptide docking and structure-based characterization of peptide binding: From knowledge to know-how. *Curr. Opin. Struct. Biol.* **23**, 894-902, (2013).

14   Yang, L. & Schepartz, A. Relationship between folding and function in a sequence-specific miniature DNA-binding protein. *Biochemistry* **44**, 7469-7478, (2005).

15   Eckel, R. *et al.* Identification of binding mechanisms in single molecule–DNA complexes. *Biophys. J.* **85**, 1968-1973, (2003).

16   Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384-396, (2014).

17   Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem* **75**, 333-366, (2006).

18   Hauser, C. a. E. *et al.* Natural tri- to hexapeptides self-assemble in water to amyloid beta-type fiber aggregates by unexpected alpha-helical intermediate structures. *Proc. Natl. Acad. Sci. USA* **108**, 1361-1366, (2011).

19   Bernstein, S. L. *et al.* Amyloid-beta protein oligomerization and the importance of tetramers and dodecamers in the aetiology of alzheimer's disease. *Nat. Chem.* **1**, 1-6, (2009).

20   Tysoe, C. *et al.* Potent human alpha-amylase inhibition by the beta-defensin-like protein helianthamide. *ACS Cent. Sci.* **2**, 154-161, (2016).

21   Reddy, K. V. R., Yedery, R. D. & Aranha, C. Antimicrobial peptides: Premises and promises. *Int. J. Antimicrob. Agents* **24**, 536-547, (2004).

22   Fjell, C. D., Hiss, J. A., Hancock, R. E. W. & Schneider, G. Designing antimicrobial peptides: Form follows function. *Nat. Rev. Drug Discov.* **11**, 37-51, (2012).

23   Tabata, A. *et al.* Development of a sortase a-mediated peptide-labeled liposome applicable to drug-delivery systems. *Anticancer Res.* **35**, 4411-4417, (2015).

24   Khafagy, E.-S. & Morishita, M. Oral biodrug delivery using cell-penetrating peptide. *Adv. Drug Del. Rev.* **64**, 531-539, (2012).

25   Saxena, R. & Nanjan, M. J. Elastin-like polypeptides and their applications in anticancer drug delivery systems: A review. *Drug Deliv.* **22**, 156-167, (2015).

26   Li, S. *et al.* Design of asymmetric peptide bilayer membranes. *J. Am. Chem. Soc.* **138**, 3579-3586, (2016).

27   Katz, C. *et al.* Studying protein-protein interactions using peptide arrays. *Chem. Soc. Rev.* **40**, 2131-2145, (2011).

28   Kaltashov, I. A. & Eyles, S. J. in *Mass spectrometry in biophysics: Conformation and dynamics of biomolecules*   45-86 (John Wiley & Sons, 2005).

29   http://www.rcsb.org/pdb/.

30   Berman, H. M. *et al.* The protein data bank. *Nucleic Acids Res.* **28**, 235-242, (2000).

31   Sherwood, D. & Cooper, J. *Crystals, X-rays, and proteins : Comprehensive protein crystallography.* (Oxford University Press, 2011).

32   Benvenuti, M. & Mangani, S. Crystallization of soluble proteins in vapor diffusion for X-ray crystallography. *Nat. Protoc.* **2**, 1633-1651, (2007).

33   Chruszcz, M. *et al.* Analysis of solvent content and oligomeric states in protein crystals—does symmetry matter? *Protein Sci.* **17**, 623-632, (2008).

34   Eyal, E., Gerzon, S., Potapov, V., Edelman, M. & Sobolev, V. The limit of accuracy of protein modeling: Influence of crystal packing on protein structure. *J. Mol. Biol.* **351**, 431-442, (2005).

35   Karle, I. L. Flexibility in peptide molecules and restraints imposed by hydrogen bonds, the AIB residue, and core inserts. *Pept. Sci.* **40**, 157-180, (1996).

36   Marraud, M. & Aubry, A. Crystal structures of peptides and modified peptides. *Pept. Sci.* **40**, 45-83, (1996).

37   Paluch, P. *et al.* Analysis of local molecular motions of aromatic sidechains in proteins by 2D and 3D fast MAS NMR spectroscopy and quantum mechanical calculations. *Phys. Chem. Chem. Phys.* **17**, 28789-28801, (2015).

38   Porcelli, F., Ramamoorthy, A., Barany, G. & Veglia, G. in *Membrane proteins: Folding, association, and design* (eds Giovanna Ghirlanda & Alessandro Senes) 159-180 (Humana Press, 2013).

39   Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964-972, (2007).

40   Wishart, D. S. & Sykes, B. D. Chemical shifts as a tool for structure determination. *Methods Enzymol.* **239**, 363-392, (1994).

41   Gronenborn, A. M. & Clore, G. M. Structures of protein complexes by multidimentional heteronuclear magnetic resonance spectroscopy. *Crit. Rev. Biochem. Mol. Biol.* **30**, 351, (1995).

42   Bechinger, B. The structure, dynamics and orientation of antimicrobial peptides in membranes by multidimensional solid-state NMR spectroscopy. *Biochim. Biophys. Acta* **1462**, 157-183, (1999).

43   Havel, T. F. An evaluation of computational strategies for use in the determination of protein-structure from distance constraints   obtained by nuclear-magnetic resonance. *Prog. Biophys. Mol. Biol.* **56**, 43-78, (1991).

44   Guntert, P. Structure calculation of biological macromolecules from NMR data. *Q. Rev. Biophys.* **31**, 145-237, (1998).

45   Micsonai, A. *et al.* Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. USA* **112**, E3095-E3103, (2015).

46   Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876-2890, (2006).

47   Bai, X.-c., McMullan, G. & Scheres, S. H. W. How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci* **40**, 49-57, (2015).

48   Binshtein, E. & Ohi, M. D. Cryo-electron microscopy and the amazing race to atomic resolution. *Biochemistry* **54**, 3133-3141, (2015).

49   Cheng, Y., Grigorieff, N., Penczek, Pawel A. & Walz, T. A primer to single-particle cryo-electron microscopy. *Cell* **161**, 438-449, (2015).

50   Mierke, D. F., Kurz, M. & Kessler, H. Peptide flexibility and calculations of an ensemble of molecules. *J. Am. Chem. Soc.* **116**, 1042-1049, (1994).

51   Mittag, T. & Forman-Kay, J. D. Atomic-level characterization of disordered protein ensembles. *Curr. Opin. Chem. Biol.* **17**, 3-14, (2007).

52   Jensen, M. R. *et al.* Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings. *Structure* **17**, 1169-1185, (2009).

53   Dunker, A. K., Obradovic, Z., Romero, P., Garner, E. C. & Brown, C. J. Intrinsic protein disorder in complete genomes. *Genome informatics* **11**, 161-171, (2000).

54   Uversky, V. N. Intrinsically disordered proteins from A to Z. *Int. J. Biochem. Cell Biol.* **43**, 1090-1103, (2011).

55   Wright, P. & Jolla, L. Breaking the protein rules. *Nature* **471**, 151-153, (2011).

56   Grossman, E., Medalia, O. & Zwerger, M. Functional architecture of the nuclear pore complex. *Annu. Rev. Biophys.* **41**, 557-584, (2012).

57   Milles, S. & Lemke, E. A. Single molecule study of the intrinsically disordered FG-repeat nucleoporin 153. *Biophys. J.* **101**, 1710-1719, (2011).

58   Canon, F. *et al.* Folding of a salivary intrinsically disordered protein upon binding to tannins. *J. Am. Chem. Soc.* **133**, 7847-7852, (2011).

59   Turoverov, K. K., Kuznetsova, I. M. & Uversky, V. N. The protein kingdom extended: Ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. *Prog. Biophys. Mol. Biol.* **102**, 73-84, (2010).

60   Uversky, V. N., Oldfield, C. J. & Dunker, A. K. Intrinsically disordered proteins in human diseases: Introducing the D2 concept. *Annu. Rev. Biophys.* **37**, 215-246, (2008).

61   Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197-208, (2005).

62   Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. & Jones, D. T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **337**, 635-645, (2004).

63   Kjaergaard, M. Can proteins be intrinsically disordered inside a membrane? *Intrinsically Disord. Proteins* **3**, 1-7, (2015).

64   Sahu, D., Bastidas, M. & Showalter, S. A. Generating NMR chemical shift assignments of intrinsically disordered proteins using carbon-detected NMR methods. *Anal. Biochem.* **449**, 17-25, (2014).

65   Jensen, M. R., Zweckstetter, M., Huang, J. R. & Blackledge, M. Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. *Chem. Rev.* **114**, 6632-6660, (2014).

66   Balasubramaniam, D. & Komives, E. Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochim. Biophys. Acta* **1834**, 1202-1209, (2012).

67   Jurneczko, E. *et al.* Intrinsic disorder in proteins: A challenge for (un)structural biology met by ion mobility-mass spectrometry. *Biochem. Soc. Trans.* **40**, 1021-1026, (2012).

68   Beveridge, R., Chappuis, Q., Macphee, C. & Barran, P. Mass spectrometry methods for intrinsically disordered proteins. *Analyst* **138**, 32-42, (2013).

69   Williams, D. M. & Pukala, T. L. Novel insights into protein misfolding diseases revealed by ion mobility-mass spectrometry. *Mass Spectrom. Rev.* **32**, 169-187, (2013).

70   Morrison, L. J. & Wysocki, V. H. Low energy CID and action IRMPD provide insights into a minor subpopulation of the gas-phase conformers of triply charged bradykinin. *Int. J. Mass spectrom.* **391**, 2-10, (2015).

71   Dian, B. C., Longarte, A. & Zwier, T. S. Conformational dynamics in a dipeptide after single-mode vibrational excitation. *Science* **296**, 2369-2373, (2002).

72   Dian, B. C., Clarkson, J. R. & Zwier, T. S. Direct measurement of energy thresholds to conformational isomerization in tryptamine. *Science* **303**, 1169-1173, (2004).

73   Armentrout, P. B. Mass spectrometry - not just a structural tool: The use of guided ion beam tandem mass spectrometry to determine thermochemistry. *J. Am. Soc. Mass. Spectrom.* **13**, 419-434, (2002).

74   Pierson, N. A. & Clemmer, D. E. An IMS–IMS threshold method for semi-quantitative determination of activation barriers: Interconversion of proline cis↔trans forms in triply protonated bradykinin. *Int. J. Mass spectrom.* **377**, 646-654, (2015).

75   Shi, L. *et al.* Ion mobility-mass spectrometry reveals the energetics of intermediates that guide polyproline folding. *J. Am. Soc. Mass. Spectrom.* **27**, 22-30, (2016).

76   Morrison, L. J. & Wysocki, V. H. Gas-phase helical peptides mimic solution-phase behavior. *J. Am. Chem. Soc.*, (2014).

77   Busch, D. J. *et al.* Intrinsically disordered proteins drive membrane curvature. *Nat. Commun.* **6**, doi:10.1038/ncomms8875, (2015).

78   Barrera, N. P., Di Bartolo, N., Booth, P. J. & Robinson, C. V. Micelles protect membrane complexes from solution to vacuum. *Science* **321**, 243-246, (2008).

79   Borysik, A. J., Hewitt, D. J. & Robinson, C. V. Detergent release prolongs the lifetime of native-like membrane protein conformations in the gas-phase. *J. Am. Chem. Soc.* **135**, 6078-6083, (2013).

80   Nagornova, N. S., Rizzo, T. R. & Boyarkin, O. V. Interplay of intra- and intermolecular H-bonding in a progressively solvated macrocyclic peptide. *Science* **336**, 320-323, (2012).

81   Servage, K. A., Silveira, J. A., Fort, K. L., Clemmer, D. E. & Russell, D. H. Water-mediated dimerization of ubiquitin ions captured by cryogenic ion mobility-mass spectrometry. *J. Phys. Chem. Lett.* **6**, 4947-4951, (2015).

82   Silveira, J. A., Servage, K. A., Gamage, C. M. & Russell, D. H. Cryogenic ion mobility-mass spectrometry captures hydrated ions produced during electrospray ionization. *J. Phys. Chem. A* **117**, 953-961, (2013).

83   Silveira, J. A. *et al.* From solution to the gas phase: Stepwise dehydration and kinetic trapping of substance p reveals the origin of peptide conformations. *J. Am. Chem. Soc.* **135**, 19147-19153, (2013).

84   Warnke, S., von Helden, G. & Pagel, K. Protein structure in the gas phase: The influence of side-chain microsolvation. *J. Am. Chem. Soc.* **135**, 1177-1180, (2013).

85   Mercier, S. R. *et al.* Microsolvation effects on the excited-state dynamics of protonated tryptophan. *J. Am. Chem. Soc.* **128**, 16938-16943, (2006).

86   Zhu, H. *et al.* Conformations and vibrational spectra of a model tripeptide: Change of secondary structure upon micro-solvation. *Phys. Chem. Chem. Phys.* **12**, 3415-3425, (2010).

87   Bechara, C. *et al.* A subset of annular lipids is linked to the flippase activity of an ABC transporter. *Nat. Chem.* **7**, 255-262, (2015).

88   Harvey, S. R. & Wysocki, V. H. Mass spectrometry: Bound in flight. *Nat. Chem.* **7**, 189-190, (2015).

89   Baldauf, C. & Rossi, M. Going clean: Structure and dynamics of peptides in the gas phase and paths to solvation. *J. Phys.: Condens. Matter* **27**, 493002, (2015).

90   Roy, T. K. *et al.* Conformational structures of a decapeptide validated by first principles calculations and cold ion spectroscopy. *ChemPhysChem* **16**, 1374-1378, (2015).

91   Zabuga, A. V. & Rizzo, T. R. Capping motif for peptide helix formation. *J. Phys. Chem. Lett.* **6**, 1504-1508, (2015).

92   Granata, D. *et al.* The inverted free energy landscape of an intrinsically disordered peptide by simulations and experiments. *Sci. Rep.* **5**, 15449, (2015).

93   Barran, P. *et al.* Is it biologically relevant to measure the structures of small peptides in the gas-phase? *Int. J. Mass spectrom.* **240**, 273-284, (2005).

94   Harvey, S. R., Macphee, C. E. & Barran, P. E. Ion mobility mass spectrometry for peptide analysis. *Methods* **54**, 454-461, (2011).

95   Nyon, M. P. *et al.* An integrative approach combining ion mobility mass spectrometry, X-ray crystallography, and nuclear magnetic resonance spectroscopy to study the conformational dynamics of alpha1-antitrypsin upon ligand binding. *Protein Sci.* **24**, 1301-1312, (2015).

96   Beveridge, R. *et al.* Relating gas phase to solution conformations: Lessons from disordered proteins. *Proteomics* **15**, 2872-2883, (2015).

97   Testa, L. *et al.* Extracting structural information from charge-state distributions of intrinsically disordered proteins by non-denaturing electrospray-ionization mass spectrometry. *Intrinsically Disord. Proteins* **1**, e25068, (2013).

98   Keppel, T. R. & Weis, D. D. Mapping residual structure in intrinsically disordered proteins at residue resolution using millisecond hydrogen/deuterium exchange and residue averaging. *J. Am. Soc. Mass. Spectrom.* **26**, 547-554, (2015).

99   Balasubramaniam, D. & Komives, E. Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochim. Biophys. Acta* **1834**, 1202-1209, (2013).

100   Kalapothakis, J. M. D. *et al.* Unusual ECD fragmentation attributed to gas-phase helix formation in a conformationally dynamic peptide. *Chem. Commun. (Cambridge, U. K.)* **50**, 198-200, (2014).

101   Pagel, K., Natan, E., Hall, Z., Fersht, A. R. & Robinson, C. V. Intrinsically disordered p53 and its complexes populate compact conformations in the gas phase. *Angew. Chem. Int. Ed.* **52**, 361-365, (2013).

102   Hudgins, R. R., Woenckhaus, J. & Jarrold, M. F. High resolution ion mobility measurements for gas phase proteins: Correlation between solution phase and gas phase conformations. *Int. J. Mass Spectrom. Ion Processes* **165**, 497-597, (1997).

103   Yamashita, M. & Fenn, J. B. Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.* **88**, 4451-4459, (1984).

104   Whitehouse, C. M., Dreyer, R. N., Yamashita, M. & Fenn, J. B. Electrospray interface for liquid chromatographs and mass spectrometers. *Anal. Chem.* **57**, 675-679, (1985).

105   Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64-71, (1989).

106   Kebarle, P. & Verkerk, U. H. Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **28**, 898-917, (2009).

107   Taylor, G. I. & McEwan, A. D. The stability of a horizontal fluid interface in a vertical electric field. *J. Fluid Mech.* **22**, 1-15, (1965).

108   Marginean, I., Parvin, L., Heffernan, L. & Vertes, A. Flexing the electrified meniscus: The birth of a jet in electrosprays. *Anal. Chem.* **76**, 4202-4207, (2004).

109   Kim, H.-H., Kim, J.-H. & Ogata, A. Time-resolved high-speed camera observation of electrospray. *J. Aerosol Sci* **42**, 249-263, (2011).

110   Rayleigh, L. On the equilibrium of liquid conducting masses charged with electricity. *Philosophical Magazine Series 5* **14**, 184-186, (1882).

111   Grimm, R. L. & Beauchamp, J. L. Evaporation and discharge dynamics of highly charged multicomponent droplets generated by electrospray ionization. *J. Phys. Chem. A* **114**, 1411-1419, (2010).

112   Grimm, R. L. & Beauchamp, J. L. Evaporation and discharge dynamics of highly charged droplets of heptane, octane, and p-xylene generated by electrospray ionization. *Anal. Chem.* **74**, 6291-6297, (2002).

113   Smith, J. N., Flagan, R. C. & Beauchamp, J. L. Droplet evaporation and discharge dynamics in electrospray ionization. *J. Phys. Chem. A* **106**, 9957-9967, (2002).

114   Juraschek, R., Dülcks, T. & Karas, M. Nanoelectrospray—more than just a minimized-flow electrospray ionization source. *J. Am. Soc. Mass. Spectrom.* **10**, 300-308, (1999).

115   Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. Unraveling the mechanism of electrospray ionization. *Anal. Chem.* **85**, 2-9, (2013).

116   Iribarne, J. V. & Thomson, B. A. On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **64**, 2287-2294, (1976).

117   Loscertales, I. G. & Fernández de la Mora, J. Experiments on the kinetics of field evaporation of small ions from droplets. *J. Chem. Phys.* **103**, 5041-5060, (1995).

118   Kebarle, P. & Tang, L. From ions in solution to ions in the gas phase - the mechanism of electrospray mass spectrometry. *Anal. Chem.* **65**, 972A-986A, (1993).

119   Fernandez de la Mora, J. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta* **406**, 93-104, (2000).

120   McAllister, R. G., Metwally, H., Sun, Y. & Konermann, L. Release of native-like gaseous proteins from electrospray droplets via the charged residue mechanism: Insights from molecular dynamics simulations. *J. Am. Chem. Soc.* **137**, 12667-12676, (2015).

121   Iavarone, A. T. & Williams, E. R. Mechanism of charging and supercharging molecules in electrospray ionization. *J. Am. Chem. Soc.* **125**, 2319-2327, (2003).

122   Yue, X., Vahidi, S. & Konermann, L. Insights into the mechanism of protein electrospray ionization from salt adduction measurements. *J. Am. Soc. Mass. Spectrom.*, 1322-1331, (2014).

123   Konermann, L., Rodriguez, A. D. & Liu, J. On the formation of highly charged gaseous ions from unfolded proteins by electrospray ionization. *Anal. Chem.* **84**, 6798-6804, (2012).

124   Li, J. *et al.* Conformational effects in protein electrospray-ionization mass spectrometry. *Mass Spectrom. Rev.* **35**, 111-122, (2015).

125   Konijnenberg, A., Butterer, A. & Sobott, F. Native ion mobility-mass spectrometry and related methods in structural biology. *Biochim. Biophys. Acta* **1834**, 1239-1256, (2012).

126   Jarrold, M. F. Unfolding, refolding and hydration of proteins in the gas phase. *Acc. Chem. Res.* **32**, 360-367, (1999).

127   Jurneczko, E. & Barran, P. E. How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. *Analyst* **136**, 20-28, (2011).

128   Wyttenbach, T. & Bowers, M. T. Structural stability from solution to the gas phase: Native solution structure of ubiquitin survives analysis in a solvent-free ion mobility-mass spectrometry environment. *J. Phys. Chem. B* **115**, 12266-12275, (2011).

129   González Flórez, A. I. *et al.* Charge-induced unzipping of isolated proteins to a defined secondary structure. *Angew. Chem. Int. Ed.* **55**, 3295-3299, (2016).

130   Pagel, K. *et al.* Gas-phase IR spectra of intact α-helical coiled coil protein complexes. *Int. J. Mass spectrom.* **283**, 161-168, (2009).

131   Ruotolo, B. T., Verbeck, G. F., Thomson, L. M., Gillig, K. J. & Russell, D. H. Observation of conserved solution-phase secondary structure in gas-phase tryptic peptides. *J. Am. Chem. Soc.* **124**, 4214-4215, (2002).

132   Li, J., Taraszka, J., Counterman, A. E. & Clemmer, D. E. Influence of solvent composition and capillary temperature on the conformations of electrosprayed ions: Unfolding of compact ubiquitin conformers from pseudonative and denatured solutions. *Int. J. Mass spectrom.* **185-187**, 37-47, (1999).

133   Loo, J. A., He, J. X. & Cody, W. L. High order structure in gas phase reflects solution structure. *J. Am. Chem. Soc.* **120**, 4542-4543, (1998).

134   Wright, P. J., Zhang, J. M. & Douglas, D. J. Conformations of gas-phase ions of ubiquitin, cytochrome c, apomyoglobin, and beta-lactoglobulin produced from two different solution conformations. *J. Am. Soc. Mass. Spectrom.* **19**, 1906-1913, (2008).

135 Shi, H. & Clemmer, D. E. Evidence for two new solution states of ubiquitin by IMS–MS analysis. *J. Phys. Chem. B* **118**, 3498-3506, (2014).

136 Chen, S. H. & Russell, D. H. How closely related are conformations of protein ions sampled by IM-MS to native solution structures? *J. Am. Soc. Mass. Spectrom.* **26**, 1433-1443, (2015).

137 Myung, S., Badman, E. R., Lee, Y. J. & Clemmer, D. E. Structural transitions of electrosprayed ubiquitin ions stored in an ion trap over 10 ms to 30 s. *J. Phys. Chem. A* **106**, 9976-9982, (2002).

138 Badman, E. R., Myung, S. & Clemmer, D. E. Evidence for unfolding and refolding of gas-phase cytochrome c ions in a paul trap. *J. Am. Soc. Mass. Spectrom.* **16**, 1493-1497, (2005).

139 Pierson, N. A., Valentine, S. J. & Clemmer, D. E. Evidence for a quasi-equilibrium distribution of states for bradykinin [m + 3h]3+ ions in the gas phase. *J. Phys. Chem. B* **114**, 7777-7783, (2010).

140 Hopper, J. T. S. & Oldham, N. J. Collision induced unfolding of protein ions in the gas phase studied by ion mobility-mass spectrometry: The effect of ligand binding on conformational stability. *J. Am. Soc. Mass. Spectrom.* **20**, 1851-1858, (2009).

141 Shi, H., Atlasevich, N., Merenbloom, S. I. & Clemmer, D. E. Solution dependence of the collisional activation of ubiquitin [m + 7h](7+) ions. *J. Am. Soc. Mass. Spectrom.* **25**, 2000-2008, (2014).

142 Fort, K. L. *et al.* From solution to the gas phase: Factors that influence kinetic trapping of substance p in the gas phase. *J. Phys. Chem. B* **118**, 14336-14344, (2014).

143 Servage, K. A., Silveira, J. A., Fort, K. L. & Russell, D. H. Cryogenic ion mobility-mass spectrometry: Tracking ion structure from solution to the gas phase. *Acc. Chem. Res.* **49**, 1421-1428, (2016).

144 Sharon, M. & Horovitz, A. Probing allosteric mechanisms using native mass spectrometry. *Curr. Opin. Struct. Biol.* **34**, 7-16, (2015).

145 Boeri Erba, E. & Petosa, C. The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes. *Protein Sci.* **24**, 1176-1192, (2015).

146 Konermann, L., Vahidi, S. & Sowole, M. A. Mass spectrometry methods for studying structure and dynamics of biological macromolecules. *Anal. Chem.* **86**, 213-232, (2014).

147 Vandermarliere, E., Stes, E., Gevaert, K. & Martens, L. Resolution of protein structure by mass spectrometry. *Mass Spectrom. Rev.*, DOI 10.1002/mas.21450, (2014).

148 Weiss, M. S., Jabs, A. & Hilgenfeld, R. Peptide bonds revisited. *Nat. Struct. Biol.* **5**, 676, (1998).

149 Cheng, H. N. & Bovey, F. A. Cis-trans equilibrium and kinetic studies of acetyl-l-proline and glycyl-l-proline. *Biopolymers* **16**, 1465-1472, (1977).

150 Vunnam, N. & Pedigo, S. Prolines in beta a-sheet of neural cadherin act as a switch to control the dynamics of the equilibrium between monomer and dimer. *Biochemistry* **50**, 6959-6965, (2011).

151 Schmidpeter, P. A. M., Koch, J. R. & Schmid, F. X. Control of protein function by prolyl isomerization. *Biochim. Biophys. Acta* **1850**, 1973-1982, (2015).

152 Reimer, U. & Fischer, G. Local structural changes caused by peptidyl-prolyl cis/trans isomerization in the native state of proteins. *Biophys. Chem.* **96**, 203-212, (2002).

153 Andreotti, A. H. Native state proline isomerization: An intrinsic molecular switch. *Biochemistry* **42**, 9515-9523, (2003).

154 Lu, K. P., Finn, G., Lee, T. H. & Nicholson, L. K. Prolyl cis-trans isomerization as a molecular timer. *Nat. Chem. Biol.* **3**, 619-629, (2007).

155 Schmid, F. X. Prolyl isomerase: Enzymatic catalysis of slow protein-folding reactions. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 123-143, (1993).

156 Wedemeyer, W. J., Welker, E. & Scheraga, H. A. Proline cis - trans isomerization and protein folding. *Biochemistry* **41**, 14637-14644, (2002).

157 Rousseau, F., Schymkowitz, J. W. H., Wilkinson, H. R. & Itzhaki, L. S. Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proc. Natl. Acad. Sci. USA* **98**, 5596-5601, (2001).

158 London, R. E., Stewart, J. M. & Cann, J. R. NMR studies of bradykinin and related peptides. *Biochem. Pharmacol.* **40**, 41-48, (1990).

159 Williamson, M. P. The structure and function of proline-rich regions in proteins. *Biochem. J* **297**, 249-260, (1994).

160 Lovell, S. C. *et al.* Structure validation by c α geometry: Φ, ϕ and c β deviation. *Proteins: Struct. Funct. Bioinform.* **50**, 437-450, (2003).

161 Pal, D. & Chakrabarti, P. Cis peptide bonds in proteins: Residues involved, their conformations, interactions and locations. *J. Mol. Biol.* **294**, 271-288, (1999).

162 Theillet, F.-X. *et al.* The alphabet of intrinsic disorder. *Intrinsically Disord. Proteins* **1**, e24360, (2013).

163 Steward, A., Adhya, S. & Clarke, J. Sequence conservation in Ig-like domains: The role of highly conserved proline residues in the fibronectin type III superfamily. *J. Mol. Biol.* **318**, 935-940, (2002).

164 Lopez-Martinez, C. *et al.* Proline cis-trans isomerization and its implications for the dimerization of analogues of cyclopeptide stylostatin 1: A combined computational and experimental study. *Phys. Chem. Chem. Phys.* **18**, 12755-12767, (2016).

165 Glover, M. S., Bellinger, E. P., Radivojac, P. & Clemmer, D. E. Penultimate proline in neuropeptides. *Anal. Chem.* **87**, 8466-8472, (2015).

166 Breci, L. A., Tabb, D. L., Yates, J. R. & Wysocki, V. H. Cleavage N-terminal to proline : Analysis of a database of peptide tandem mass spectra. *Anal. Chem.* **75**, 1963-1971, (2003).

167 Bleiholder, C., Suhai, S., Harrison, A. G. & Paizs, B. Towards understanding the tandem mass spectra of protonated oligopeptides. 2: The proline effect in collision-induced dissociation of protonated Ala-Ala-Xxx-Pro-Ala (Xxx = Ala, Ser, Leu, Val, Phe, and Trp). *J. Am. Soc. Mass. Spectrom.* **22**, 1032-1039, (2011).

168 Kim, T.-Y., Valentine, S. J., Clemmer, D. E. & Reilly, J. P. Gas-phase conformation-specific photofragmentation of proline-containing peptide ions. *J. Am. Soc. Mass. Spectrom.* **21**, 1455-1465, (2010).

169 Warnke, S., Baldauf, C., Bowers, M. T., Pagel, K. & von Helden, G. Photodissociation of conformer-selected ubiquitin ions reveals site-specific cis/trans isomerization of proline peptide bonds. *J. Am. Chem. Soc.* **136**, 10308-10314, (2014).

170 Gucinski, A. C., Chamot-Rooke, J., Steinmetz, V., Somogyi, A. & Wysocki, V. H. Influence of N−terminal residue composition on the structure of proline-containing b2+ ions. *J. Phys. Chem. A* **117**, 1291-1298, (2013).

171 Martens, J. K., Grzetic, J., Berden, G. & Oomens, J. Gas-phase conformations of small polyprolines and their fragment ions by IRMPD spectroscopy. *Int. J. Mass spectrom.* **377**, 179-187, (2014).

172 Bernier, M., Chamot-Rooke, J. & Wysocki, V. R vs. S fluoroproline ring substitution: Trans/cis effects on the formation of b2 ions in gas-phase peptide fragmentation. *Phys. Chem. Chem. Phys.* **18**, 2202, (2016).

173   Pierson, N. A., Chen, L., Russell, D. H. & Clemmer, D. E. Cis-trans isomerizations of proline residues are key to bradykinin conformations. *J. Am. Chem. Soc.* **135**, 3186-3192, (2013).

174   Glover, M. S. *et al.* On the split personality of penultimate proline. *J. Am. Soc. Mass. Spectrom.* **26**, 444-452, (2015).

175   Counterman, A. E. & Clemmer, D. E. Cis-trans signatures of proline-containing tryptic peptides in the gas phase. *Anal. Chem.* **74**, 1946-1951, (2002).

176   Fuller, D. R. *et al.* Cis→trans isomerization of Pro7 in oxytocin regulates Zn2+ binding. *J. Am. Soc. Mass. Spectrom.* **27**, 1376-1382, (2016).

177   Regoli, D. & Barabe, J. Pharmacology of bradykinin and related kinins. *Pharmacol. Rev.* **32**, 1-46, (1980).

178   Takano, M. & Matsuyama, S. Intracellular and nuclear bradykinin B2 receptors. *Eur. J. Pharmacol.* **732**, 169-172, (2014).

179   Denys, L., Bothner-By, A. A., Fisher, G. H. & Ryan, J. W. Conformational diversity of bradykinin in aqueous solution. *Biochemistry* **21**, 6531-6536, (1982).

180   Young, J. K. & Hicks, R. P. NMR and molecular modeling investigations of the neuropeptide bradykinin in three different solvent systems: DMSO, 9:1 dioxane/water, and in the presence of 7.4 mM lyso phosphatidylcholine micelles. *Biopolymers* **34**, 611-623, (1994).

181   Gaggelli, E., Amelio, N. D., Maccotta, A. & Valensin, G. Calcium-binding properties and molecular organization of bradykinin. *Eur. J. Biochem.* **262**, 268-276, (1999).

182   Richard, T., Vergé, S., Berké, B., Vercauteren, J. & Monti, J. P. NMR and simulated annealing investigations of bradykinin in presence of polyphenols. *J. Biomol. Struct. Dyn.* **18**, 627-637, (2001).

183   Bonechi, C., Ristori, S., Martini, G., Martini, S. & Rossi, C. Study of bradykinin conformation in the presence of model membrane by nuclear magnetic resonance and molecular modelling. *Biochim. Biophys. Acta* **1788**, 708-716, (2009).

184   Ottleben, H. *et al.* An NMR study of the interaction of 15N-labelled bradykinin with an antibody mimic of the bradykinin B2 receptor. *Eur. J. Biochem.* **244**, 471-478, (1997).

185   Lopez, J. J. *et al.* The structure of the neuropeptide bradykinin bound to the human G-protein coupled receptor bradykinin B2 as determined by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed.* **47**, 1668-1671, (2008).

186   Richard, T., Delaunay, J. C., Mérillon, J. M. & Monti, J. P. Is the C-terminal region of bradykinin the binding site of polyphenols? *J. Biomol. Struct. Dyn.* **21**, 379-385, (2003).

187   Chatterjee, C. & Mukhopadhyay, C. Interaction and structural study of kinin peptide bradykinin and ganglioside monosialylated 1 micelle. *Biopolymers* **78**, 197-205, (2005).

188   Cann, J. R., Liu, X., Stewart, J. M., Gera, L. & Kotovych, G. A CD and an NMR study of multiple bradykinin conformations in aqueous trifluoroethanol solutions. *Biopolymers* **34**, 869-878, (1994).

189   Shvartsburg, A. & Smith, R. D. Fundamentals of traveling wave ion mobility spectrometry. *Anal. Chem.* **80**, 9689-9699, (2008).

190   Wyttenbach, T., Helden, G. V. & Bowers, M. T. Gas-phase conformation of biological molecules : Bradykinin. *J. Am. Chem. Soc.* **118**, 8355-8364, (1996).

191   Purves, R. W., Barnett, D. A., Ells, B. & Guevremont, R. Gas-phase conformers of the [m + 2h](2+) ion of bradykinin investigated by combining high-field asymmetric waveform ion mobility spectrometry, hydrogen/deuterium exchange, and energy-loss measurements. *Rapid Commun. Mass Spectrom.* **15**, 1453-1456, (2001).

192   Rodriquez, C. F. *et al.* Gaseous bradykinin and its singly, doubly, and triply protonated forms: A first-principles study. *J. Phys. Chem. B* **110**, 7528-7537, (2006).

193   Counterman, A. E. *et al.* High-order structure and dissociation of gaseous peptide aggregates that are hidden in mass spectra. *J. Am. Soc. Mass. Spectrom.* **9**, 743-759, (1998).

194   Guo, Y., Wang, J., Javahery, G., Thomson, B. A. & Siu, K. W. M. Ion mobility spectrometer with radial collisional focusing. *Anal. Chem.* **77**, 266-275, (2005).

195   Papadopoulos, G., Svendsen, A., Boyarkin, O. V. & Rizzo, T. R. Conformational distribution of bradykinin [bk + 2 h](2+) revealed by cold ion spectroscopy coupled with FAIMS. *J. Am. Soc. Mass. Spectrom.* **23**, 1173-1181, (2012).

196   Pierson, N. A., Chen, L. X., Valentine, S. J., Russell, D. H. & Clemmer, D. E. Number of solution states of bradykinin from ion mobility and mass spectrometry measurements. *J. Am. Chem. Soc.* **133**, 13810-13813, (2011).

197   Compagnon, I., Oomens, J., Meijer, G. & von Helden, G. Mid-infrared spectroscopy of protected peptides in the gas phase: A probe of the backbone conformation. *J. Am. Chem. Soc.* **128**, 3592-3597, (2006).

198   Murphey, L. J., Hachey, D. L., Oates, J. A., Morrow, J. D. & Brown, N. J. Metabolism of bradykinin in vivo in humans: Identification of BK1-5 as a stable plasma peptide metabolite. *J. Pharmacol. Exp. Ther.* **294**, 263-269, (2000).

199   Morinelli, T. A., Webb, J. G., Jaffa, A. A., Privitera, P. J. & Margolius, H. S. A metabolic fragment of bradykinin, Arg-Pro-Pro-Gly-Phe, protects against the deleterious effects of lipopolysaccharide in rats. *J. Pharmacol. Exp. Ther.* **296**, 71-76, (2001).

200   Hasan, A. A. K. *et al.* Mechanisms of Arg-Pro-Pro-Gly-Phe inhibition of thrombin. *Am. J. Physiol. Heart Circ. Physiol.* **285**, 183-193, (2003).

201   Nieman, M. T. *et al.* Mapping the interaction of bradykinin 1–5 with the exodomain of human protease activated receptor 4. *FEBS Lett.* **579**, 25-29, (2005).

202   Sawyer, H. a. *et al.* The structure of gas-phase bradykinin fragment 1-5 (RPPGF) ions: An ion mobility spectrometry and H/D exchange ion-molecule reaction chemistry study. *J. Am. Soc. Mass. Spectrom.* **16**, 893-905, (2005).

203   Goldstein, M., Zmiri, L., Segev, E., Wyttenbach, T. & Gerber, R. B. An atomistic structure of ubiquitin +13 relevant in mass spectrometry: Theoretical prediction and comparison with experimental cross sections. *Int. J. Mass spectrom.* **367**, 10-15, (2014).

204   Cismesia, A. P., Bailey, L. S., Bell, M. R., Tesler, L. F. & Polfer, N. C. Making mass spectrometry see the light: The promises and challenges of cryogenic infrared ion spectroscopy as a bioanalytical technique. *J. Am. Soc. Mass. Spectrom.* **27**, 757-766, (2016).

205   Polfer, N. C. Infrared multiple photon dissociation spectroscopy of trapped ions. *Chem. Soc. Rev.* **40**, 2211-2221, (2011).

206   Oomens, J., Steill, J. D. & Redlich, B. Gas-phase IR spectroscopy of deprotonated amino acids. *J. Am. Chem. Soc.* **131**, 4310-4319, (2009).

207   Gonzalez Florez, A. I., Ahn, D. S., Gewinner, S., Schollkopf, W. & von Helden, G. IR spectroscopy of protonated leu-enkephalin and its 18-crown-6 complex embedded in helium droplets. *Phys. Chem. Chem. Phys.* **17**, 21902-21911, (2015).

208   Oomens, J., Sartakov, B. G., Meijer, G. & von Helden, G. Gas-phase infrared multiple photon dissociation spectroscopy of mass-selected molecular ions. *Int. J. Mass spectrom.* **254**, 1-19, (2006).

209   Kamrath, M. Z. *et al.* Vibrational characterization of simple peptides using cryogenic infrared photodissociation of H2-tagged, mass-selected ions. *J. Am. Chem. Soc.* **133**, 6440-6448, (2011).

210   Oomens, J. *et al.* Charge-state resolved mid-infrared spectroscopy of a gas-phase protein. *Phys. Chem. Chem. Phys.* **7**, 1345-1348, (2005).

211   Rizzo, T. R., Stearns, J. A. & Boyarkin, O. V. Spectroscopic studies of cold, gas-phase biomolecular ions. *Int. Rev. Phys. Chem.* **28**, 481-515, (2009).

212   Boyarkin, O. V., Mercier, S. R., Kamariotis, A. & Rizzo, T. R. Electronic spectroscopy of cold, protonated tryptophan and tyrosine. *J. Am. Chem. Soc* **128**, 2816-2817, (2006).

213   Stearns, J., Seaiby, C., Boyarkin, O. V. & Rizzo, T. R. Spectroscopy and conformational preferences of gas-phase helices. *Phys. Chem. Chem. Phys.* **11**, 125-132, (2009).

214   Nagornova, N. S. *et al.* Cold-ion spectroscopy reveals the intrinsic structure of a decapeptide. *Angew. Chem. Int. Ed.* **50**, 5383-5386, (2011).

215   Burke, N. L., Redwine, J. G., Dean, J. C., McLuckey, S. A. & Zwier, T. S. UV and IR spectroscopy of cold protonated leucine enkephalin. *Int. J. Mass spectrom.* **378**, 196-205, (2015).

216   Nagornova, N. S., Rizzo, T. R. & Boyarkin, O. V. Exploring the mechanism of IR-UV double-resonance for quantitative spectroscopy of protonated polypeptides and proteins. *Angew. Chem. Int. Ed.* **52**, 6002-6005, (2013).

217   Warnke, S., von Helden, G. & Pagel, K. Analyzing the higher order structure of proteins with conformer-selective ultraviolet photodissociation. *Proteomics* **15**, 2804-2812, (2015).

218   Masson, A. *et al.* Infrared spectroscopy of mobility-selected H-Gly-Pro-Gly-Gly (GPGG). *J. Am. Soc. Mass. Spectrom.* **26**, 1444-1454, (2015).

219   Warnke, S. *et al.* Protomers of benzocaine: Solvent and permittivity dependence. *J. Am. Chem. Soc.* **137**, 4236-4242, (2015).

220   Svendsen, A., Lorenz, U. J., Boyarkin, O. V. & Rizzo, T. R. A new tandem mass spectrometer for photofragment spectroscopy of cold, gas-phase molecular ions. *Rev. Sci. Instrum.* **81**, 073107, (2010).

221   Kim, T. *et al.* Design and implementation of a new electrodynamic ion funnel. *Anal. Chem.* **72**, 2247-2255, (2000).

222   Kelly, R. T., Tolmachev, A. V., Page, J. S., Tang, K. & Smith, R. D. The ion funnel: Theory, implementations, and applications. *Mass Spectrom. Rev.* **29**, 294-312, (2010).

223   Douglas, D. J. Applications of collision dynamics in quadrupole mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **9**, 101-113, (1998).

224   Boyarkin, O. V. & Kopysov, V. Cryogenically cooled octupole ion trap for spectroscopy of biomolecular ions. *Rev. Sci. Instrum.* **85**, 033105, (2014).

225   Aseev, O. *Spectroscopic studies of peptide fragments produced by collision-induced dissociation* PhD thesis, EPFL, (2014).

226   Zabuga, A. V., Kamrath, M. Z., Boyarkin, O. V. & Rizzo, T. R. Fragmentation mechanism of UV-excited peptides in the gas phase. *J. Chem. Phys.* **141**, 154309-154309, (2014).

227   Lorenz, U. J. & Rizzo, T. R. Structural melting of an amino acid dimer upon intersystem crossing. *J. Am. Chem. Soc.* **136**, 14974-14980, (2014).

228   Guidi, M., Lorenz, U. J., Papadopoulos, G., Boyarkin, O. V. & Rizzo, T. R. Spectroscopy of protonated peptides assisted by infrared multiple photon excitation. *J. Phys. Chem. A* **113**, 797-799, (2009).

229   Shvartsburg, A. A. *Differential ion mobility spectrometry: Nonlinear ion transport and fundamentals of FAIMS*. 1st edn, (CRC Press, 2009).

230   Purves, R. W., Guevremont, R., Day, S., Pipich, C. W. & Matyjaszczyk, M. S. Mass spectrometric characterization of a high-field asymmetric waveform ion mobility spectrometer. *Rev. Sci. Instrum.* **69**, 4094-4105, (1998).

231   Krylova, N., Krylov, E., Eiceman, G. A. & Stone, J. A. Effect of moisture on the field dependence of mobility for gas-phase ions of organophosphorus compounds at atmospheric pressure with field asymmetric ion mobility spectrometry. *J. Phys. Chem. A* **107**, 3648-3654, (2003).

232   Nazarov, E. G., Coy, S. L., Krylov, E. V., Miller, R. A. & Eiceman, G. A. Pressure effects in differential mobility spectrometry. *Anal. Chem.* **78**, 7697-7706, (2006).

233   Campbell, J. L., Zhu, M. & Hopkins, W. S. Ion-molecule clustering in differential mobility spectrometry: Lessons learned from tetraalkylammonium cations and their isomers. *J. Am. Soc. Mass. Spectrom.* **25**, 1583-1591, (2014).

234   Shvartsburg, A. A. *et al.* Field asymmetric waveform ion mobility spectrometry studies of proteins: Dipole alignment in ion mobility spectrometry? *J. Phys. Chem. B* **110**, 21966-21980, (2006).

235   Shvartsburg, A. A., Noskov, S. Y., Purves, R. W. & Smith, R. D. Pendular proteins in gases and new avenues for characterization of macromolecules by ion mobility spectrometry. *Proc. Natl. Acad. Sci. USA* **106**, 6495-6500, (2009).

236   Buryakov, I. A., Krylov, E. V., Nazarov, E. G. & Rasulev, U. K. A new method of separation of multi-atomic ions by mobility at atmospheric pressure using a high-frequency amplitude-asymmetric strong electric field. *Int. J. Mass Spectrom. Ion Processes* **128**, 143-148, (1993).

237   Krylov, E. V. A method of reducing diffusion losses in a drift spectrometer. *Tech. Phys.* **44**, 113-116, (1999).

238   Guevremont, R., Barnett, D. A., Purves, R. W. & Vandermey, J. Analysis of a tryptic digest of pig hemoglobin using ESI-FAIMS-MS. *Anal. Chem.* **72**, 4577-4584, (2000).

239   Barnett, D. A., Purves, R. W., Ells, B. & Guevremont, R. Separation of o-, m- and p-phthalic acids by high-field asymmetric waveform ion mobility spectrometry (FAIMS) using mixed carrier gases. *J. Mass Spectrom.* **35**, 976-980, (2000).

240   Guevremont, R., Ding, L., Ells, B., Barnett, D. a. & Purves, R. W. Atmospheric pressure ion trapping in a tandem FAIMS-FAIMS coupled to a TOFMS: Studies with electrospray generated gramicidin s ions. *J. Am. Soc. Mass. Spectrom.* **12**, 1320-1330, (2001).

241   Barnett, D., Belford, M., Dunyach, J. J. & Purves, R. W. Characterization of a temperature-controlled FAIMS system. *J. Am. Soc. Mass. Spectrom.* **18**, 1653-1663, (2007).

242   Kapron, J. T., Jemal, M., Duncan, G., Kolakowski, B. & Purves, R. Removal of metabolite interference during liquid chromatography/tandem mass spectrometry using high-field asymmetric waveform ion mobility spectrometry. *Rapid Commun. Mass Spectrom.* **19**, 1979-1983, (2005).

243   Shvartsburg, A. & Smith, R. D. Optimum waveforms for differential ion mobility spectrometry (FAIMS). *J. Am. Soc. Mass. Spectrom.* **19**, 1286-1295, (2008).

244   Viehland, L. A., Guevremontb, R., Purves, R. W. & Barnett, D. A. Comparison of high-field ion mobility obtained from drift tubes and a FAIMS apparatus. *Int. J. Mass spectrom.* **197**, 123-130, (2000).

245   Guevremont, R., Barnett, D. A., Purves, R. W. & Viehland, L. A. Calculation of ion mobilities from electrospray ionization high-field asymmetric waveform ion mobility spectrometry mass spectrometry. *J. Chem. Phys.* **114**, 10270-10277, (2001).

246   Prieto, M. *et al.* Comparison of rectangular and bisinusoidal waveforms in a miniature planar high-field asymmetric waveform ion mobility spectrometer. *Anal. Chem.* **83**, 9237-9243, (2011).

247   McCracken, I. Automated optimization of asymmetric waveform generator LC tuning electronics. United State patent US 7058528 B2 (2006).

248   Shvartsburg, A. A., Tang, K. & Smith, R. D. Modeling the resolution and sensitivity of FAIMS analyses. *J. Am. Soc. Mass. Spectrom.* **15**, 1487-1498, (2004).

249   Shvartsburg, A. A., Tang, K. & Smith, R. D. Optimization of the design and operation of FAIMS analyzers. *J. Am. Soc. Mass. Spectrom.* **16**, 2-12, (2005).

250   Shvartsburg, A. & Smith, R. D. Ultrahigh-resolution differential ion mobility spectrometry using extended separation times. *Anal. Chem.* **83**, 23-29, (2011).

251   Purves, R. W. *et al.* Using gas modifiers to significantly improve sensitivity and selectivity in a cylindrical FAIMS device. *J. Am. Soc. Mass. Spectrom.* **25**, 1274-1284, (2014).

252   Isenberg, S. L., Armistead, P. M. & Glish, G. L. Optimization of peptide separations by differential ion mobility spectrometry. *J. Am. Soc. Mass. Spectrom.* **25**, 1592-1599, (2014).

253   Shvartsburg, A. A. & Smith, R. D. Separation of protein conformers by differential ion mobility in hydrogen-rich gases. *Anal. Chem.* **85**, 6967-6973, (2013).

254   Shvartsburg, A. A., Danielson, W. F. & Smith, R. D. High-resolution differential ion mobility separations using helium-rich gases. *Int. J. Mass spectrom.* **82**, 2456-2462, (2010).

255   Revercomb, H. E. & Mason, E. A. Theory of plasma chromatography/gaseous electrophoresis. Review. *Anal. Chem.* **47**, 970-983, (1975).

256   von Helden, G., Hsu, M.-T., Gotts, N. & Bowers, M. T. Carbon cluster cations with up to 84 atoms: Structures, formation mechanism, and reactivity. *J. Phys. Chem.* **97**, 8182-8192, (1993).

257   Clemmer, D. E. & Jarrold, M. F. Ion mobility measurements and their applications to clusters and biomolecules. *J. Mass Spectrom.* **32**, 577-592, (1997).

258   McLean, J. A., Ruotolo, B. T., Gillig, K. J. & Russell, D. H. Ion mobility–mass spectrometry: A new paradigm for proteomics. *Int. J. Mass spectrom.* **240**, 301-315, (2005).

259   Kanu, A. B., Dwivedi, P., Tam, M., Matz, L. & Hill Jr., H. H. Ion mobility – mass spectrometry. *J. Mass Spectrom.* **43**, 1-22, (2008).

260   May, J. C. & McLean, J. Ion mobility-mass spectrometry: Time-dispersive instrumentation. *Anal. Chem.* **87**, 1422-1436, (2015).

261 Merenbloom, S. I., Glaskin, R. S., Henson, Z. B. & Clemmer, D. E. High-resolution ion cyclotron mobility spectrometry. *Anal. Chem.* **81**, 1482-1487, (2009).

262 Maurer, M. M., Donohoe, G. C. & Valentine, S. J. Advances in ion mobility-mass spectrometry instrumentation and techniques for characterizing structural heterogeneity. *Analyst* **140**, 6782-6798, (2015).

263 Ewing, M. A., Glover, M. S. & Clemmer, D. E. Hybrid ion mobility and mass spectrometry as a separation tool. *J. Chromatogr. A* **1439**, 3-25, (2016).

264 Scarff, C. A., Thalassinos, K., Hilton, G. R. & Scrivens, J. H. Travelling wave ion mobility mass spectrometry studies of protein structure : Biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements. *Rapid Commun. Mass Spectrom.* **22**, 3297-3304, (2008).

265 Pringle, S. D. *et al.* An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int. J. Mass spectrom.* **261**, 1-12, (2007).

266 Giles, K. *et al.* Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun. Mass Spectrom.* **18**, 2401-2414, (2004).

267 Ruotolo, B. T., Benesch, J. L., Sandercock, A. M., Hyung, S. J. & Robinson, C. V. Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protoc.* **3**, 1139-1152, (2008).

268 Smith, D. P. *et al.* Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. *Eur. J. Mass Spectrom.* **15**, 113-130, (2009).

269 Bush, M. F. *et al.* Collision cross sections of proteins and their complexes: A calibration framework and database for gas-phase structural biology. *Anal. Chem.* **82**, (2010).

270 Sun, Y., Vahidi, S., Sowole, M. A. & Konermann, L. Protein structural studies by traveling wave ion mobility spectrometry: A critical look at electrospray sources and calibration issues. *J. Am. Soc. Mass. Spectrom.* **27**, 31-40, (2016).

271 Hamid, A. M. *et al.* Characterization of traveling wave ion mobility separations in structures for lossless ion manipulations. *Anal. Chem.* **87**, 11301-11308, (2015).

272 Zhang, X. *et al.* Ion trapping, storage, and ejection in structures for lossless ion manipulations. *Anal. Chem.* **87**, 6010-6016, (2015).

273   Webb, I. K. *et al.* Experimental evaluation and optimization of structures for lossless ion manipulations for ion mobility spectrometry with time-of-flight mass spectrometry. *Anal. Chem.* **86**, 9169-9176, (2014).

274   Ian Webb *et al.* in *ASMS*   279769 (San Antonio, TX, USA, 2016).

275   Silveira, J. A., Ridgeway, M. E. & Park, M. A. High resolution trapped ion mobility spectrometery of peptides. *Anal. Chem.* **86**, 5624-5627, (2014).

276   Ridgeway, M. E., Silveira, J. A., Meier, J. E. & Park, M. A. Microheterogeneity within conformational states of ubiquitin revealed by high resolution trapped ion mobility spectrometry. *Analyst* **140**, 6964-6972, (2015).

277   May, J. C. & Russell, D. H. A mass-selective variable-temperature drift tube ion mobility-mass spectrometer for temperature dependent ion mobility studies. *J. Am. Soc. Mass. Spectrom.* **22**, 1134-1145, (2011).

278   Servage, K. A. *et al.* Unfolding of hydrated alkyl diammonium cations revealed by cryogenic ion mobility-mass spectrometry. *J. Am. Chem. Soc.* **137**, 8916-8919, (2015).

279   Vidal-de-Miguel, G., Macía, M. & Cuevas, J. Transversal modulation ion mobility spectrometry (TM-IMS), a new mobility filter overcoming turbulence related limitations. *Anal. Chem.* **84**, 7831-7837, (2012).

280   Meyer, N. A., Root, K., Zenobi, R. & Vidal-de-Miguel, G. Gas-phase dopant-induced conformational changes monitored with transversal modulation ion mobility spectrometry. *Anal. Chem.* **88**, 2033-2040, (2016).

281   Koeniger, S. L. *et al.* An IMS-IMS analogue of MS-MS. *Anal. Chem.* **78**, 4161-4174, (2006).

282   Tang, K. *et al.* High-sensitivity ion mobility spectrometry/mass spectrometry using electrodynamic ion funnel interfaces. *Anal. Chem.* **77**, 3330-3339, (2005).

283   Hofmann, J. & Pagel, K.   (http://www.fhi-berlin.mpg.de/mp/pagel).

284   Lorenz, U. *Spectroscopy of cold , biomolecular ions : Instrumentation and application* PhD thesis, EPFL, (2011).

285   Debatin, M. *et al.* Planar multipole ion trap. *Phys. Rev. A* **77**, 033422, (2008).

286   Marmillod, P., Antonioni, S. & Lorenz, U. J. A radio frequency/high voltage pulse generator for the operation of a planar multipole ion trap/time-of-flight mass spectrometer. *Rev. Sci. Instrum.* **84**, 044707, (2013).

287   Lorenz, U. J. & Rizzo, T. R. Planar multipole ion trap/time-of-flight mass spectrometer. *Anal. Chem.* **83**, 7895-7901, (2011).

288   Wiley, W. C. & McLaren, I. H. Time‐of‐flight mass spectrometer with improved resolution. *Rev. Sci. Instrum.* **26**, 1150-1157, (1955).

289   Okumura, M., Yeh, L. I. & Lee, Y. T. Infrared spectroscopy of the cluster ions H3+·(H2)n. *J. Chem. Phys.* **88**, 79-91, (1988).

290   Okumura, M., Yeh, L. I., Myers, J. D. & Lee, Y. T. Infrared spectra of the solvated hydronium ion: Vibrational predissociation spectroscopy of mass-selected H3O+.(H2O)n.(H2)m. *J. Phys. Chem.* **94**, 3416-3427, (1990).

291   Kamrath, M. Z., Relph, R. A., Guasco, T. L., Leavitt, C. M. & Johnson, M. A. Vibrational predissociation spectroscopy of the H2-tagged mono- and dicarboxylate anions of dodecanedioic acid. *Int. J. Mass spectrom.* **300**, 91-98, (2011).

292   Goebbert, D. J., Wende, T., Bergmann, R., Meijer, G. & Asmis, K. R. Messenger-tagging electrosprayed ions: Vibrational spectroscopy of suberate dianions. *J. Phys. Chem. A* **113**, 5874-5880, (2009).

293   Brümmer, M., Kaposta, C., Santambrogio, G. & Asmis, K. R. Formation and photodepletion of cluster ion–messenger atom complexes in a cold ion trap: Infrared spectroscopy of VO+,VO2+, and VO3+. *J. Chem. Phys.* **119**, 12700-12703, (2003).

294   Jiang, L., Wende, T., Bergmann, R., Meijer, G. & Asmis, K. R. Gas-phase vibrational spectroscopy of microhydrated magnesium nitrate ions [MgNO3(H2O)1−4]+. *J. Am. Chem. Soc.* **132**, 7398-7404, (2010).

295   Myshakin, E. M., Jordan, K. D., Robertson, W. H., Weddle, G. H. & Johnson, M. A. Dominant structural motifs of NO−·(H2O)n complexes: Infrared spectroscopic and ab initio studies. *J. Chem. Phys.* **118**, 4945-4953, (2003).

296   Roscioli, J. R. & Johnson, M. A. Isomer-specific spectroscopy of the (H2O)8− cluster anion in the intramolecular bending region by selective photodepletion of the more weakly electron binding species (isomer II). *J. Chem. Phys.* **126**, 024307, (2007).

297   Duncan, M. A. Frontiers in the spectroscopy of mass-selected molecular ions. *Int. J. Mass spectrom.* **200**, 545-569, (2000).

298   Johnson, C. J. *et al.* Communication: He-tagged vibrational spectra of the SarGlyH+ and H+(H2O)2,3 ions: Quantifying tag effects in cryogenic ion vibrational predissociation (CIVP) spectroscopy. *J. Chem. Phys.* **140**, 221101, (2014).

299   Masson, A., Williams, E. R. & Rizzo, T. R. Molecular hydrogen messengers can lead to structural infidelity: A cautionary tale of protonated glycine. *J. Chem. Phys.* **143**, 104313, (2015).

300   http://mestrelab.com/software/mnova/nmr/.

301   Pappu, R. V., Hart, R. K. & Ponder, J. W. Analysis and application of potential energy smoothing and search methods for global optimization. *J. Phys. Chem. B* **102**, 9725-9742, (1998).

302   Kaminski, G. A. & Friesner, R. A. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **105**, 6474-6487, (2001).

303   Van Der Spoel, D. *et al.* Gromacs: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701-1718, (2005).

304   Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435-447, (2008).

305   Wyttenbach, T., von Helden, G., Batka Jr., J. J., Carlat, D. & Bowers, M. T. Effect of the long-range potential on ion mobility measurements. *J. Am. Soc. Mass. Spectrom.* **8**, 275-282, (1997).

306   D'Atri, V., Porrini, M., Rosu, F. & Gabelica, V. Linking molecular models with ion mobility experiments. Illustration with a rigid nucleic acid structure. *J. Mass Spectrom.* **50**, 711-726, (2015).

307   http://www.indiana.edu/~nano/software/.

308   https://labs.chem.ucsb.edu/bowers/michael/theory_analysis/cross-sections/sigma.shtml.

309   Marklund, Erik G., Degiacomi, Matteo T., Robinson, Carol V., Baldwin, Andrew J. & Benesch, Justin L. P. Collision cross sections for structural proteomics. *Structure* **23**, 791-799, (2015).

310   Wyttenbach, T., Bleiholder, C. & Bowers, M. T. Factors contributing to the collision cross section of polyatomic ions in the kilodalton to gigadalton range: Application to ion mobility measurements. *Anal. Chem.* **85**, 2191-2199, (2013).

311   Anderson, S. E., Bleiholder, C., Brocker, E. R., Stang, P. J. & Bowers, M. T. A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (III): Application to supramolecular coordination-driven assemblies with complex shapes. *Int. J. Mass spectrom.* **330-332**, 78-84, (2012).

312   Bleiholder, C., Contreras, S., Do, T. D. & Bowers, M. T. A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (II). Model parameterization and definition of empirical shape factors for proteins. *Int. J. Mass spectrom.* **345-347**, 89-96, (2013).

313   Bleiholder, C., Johnson, N. R., Contreras, S., Wyttenbach, T. & Bowers, M. T. Molecular structures and ion mobility cross sections: Analysis of the effects of He and N2 buffer gas. *Anal. Chem.* **87**, 7196-7203, (2015).

314   Shvartsburg, A. A. & Jarrold, M. F. An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.* **261**, 86-91, (1996).

315   Shvartsburg, A. A., Mashkevich, S. V., Baker, E. S. & Smith, R. D. Optimization of algorithms for ion mobility calculations. *J. Phys. Chem. A* **111**, 2002-2010, (2007).

316   Kim, H. *et al.* Experimental and theoretical investigation into the correlation between mass and ion mobility for choline and other ammonium cations in N2. *Anal. Chem.* **80**, 1928-1936, (2008).

317   Larriba, C. & Hogan, C. J. Ion mobilities in diatomic gases: Measurement versus prediction with non-specular scattering models. *J. Phys. Chem. A* **117**, 3887-3901, (2013).

318   Mesleh, M. F., Hunter, J. M., Shvartsburg, A. A., Schatz, G. C. & Jarrold, M. F. Structural information from ion mobility measurements: Effects of the long-range potential. *J. Phys. Chem.* **100**, 16082-16086, (1996).

319   Blum, V. *et al.* Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **180**, 2175-2196, (2009).

320   Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865-3868, (1996).

321   DiStasio, R. A., von Lilienfeld, O. A. & Tkatchenko, A. Collective many-body van der Waals interactions in molecular systems. *Proc. Natl. Acad. Sci. USA* **109**, 14791-14795, (2012).

322   Tkatchenko, A., Distasio, R. A., Car, R. & Scheffler, M. Accurate and efficient method for many-body van der Waals interactions. *Phys. Rev. Lett.* **108**, 1-5, (2012).

323   Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **110**, 6158-6158, (1999).

324   Rossi, M., Chutia, S., Scheffler, M. & Blum, V. Validation challenge of density-functional theory for peptides-example of Ac-Phe-Ala5-LysH(+). *J. Phys. Chem. A* **118**, 7349-7359, (2014).

325   Schubert, F. *et al.* Native like helices in a specially designed beta peptide in the gas phase. *Phys. Chem. Chem. Phys.* **17**, 5376-5385, (2015).

326   Schubert, F. *et al.* Exploring the conformational preferences of 20-residue peptides in isolation: Ac-Ala19-Lys + H+ vs. Ac-Lys-Ala19+H+ and the current reach of DFT. *Phys. Chem. Chem. Phys.* **17**, 7373-7385, (2015).

327   Voronina, L. *et al.* Conformations of prolyl–peptide bonds in the bradykinin 1–5 fragment in solution and in the gas phase. *J. Am. Chem. Soc.* **138**, 9224-9233, (2016).

328   Uversky, V. N. & Dunker, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta* **1804**, 1231-1264, (2010).

329   Buchberger, A., Yu, Q. & Li, L. Advances in mass spectrometric tools for probing neuropeptides. *Annu. Rev. Anal. Chem.* **8**, 485-509, (2015).

330   Schennach, M. & Breuker, K. Probing protein structure and folding in the gas phase by electron capture dissociation. *J. Am. Soc. Mass. Spectrom.* **26**, 1059-1067, (2015).

331   Wyttenbach, T., Pierson, N. A., Clemmer, D. E. & Bowers, M. T. Ion mobility analysis of molecular dynamics. *Annu. Rev. Phys. Chem.* **65**, 175-196, (2014).

332   Stedwell, C. N., Galindo, J. F., Roitberg, A. E. & Polfer, N. C. Structures of biomolecular ions in the gas phase probed by infrared light sources. *Annu. Rev. Anal. Chem.* **6**, 267-285, (2013).

333   van der Spoel, D., Marklund, E. G., Larsson, D. S. D. & Caleman, C. Proteins, lipids, and water in the gas phase. *Macromol. Biosci.* **11**, 50-59, (2011).

334   Breuker, K. & McLafferty, F. W. Stepwise evolution of protein native structure with electrospray into the gas phase. *Proc. Natl. Acad. Sci. USA* **105**, 18145-18152, (2008).

335   Breuker, K., Brüschweiler, S. & Tollinger, M. Electrostatic stabilization of a native protein structure in the gas phase. *Angew. Chem. Int. Ed.* **50**, 873-877, (2011).

336   Shelimov, K. B. & Jarrold, M. F. Conformations, unfolding, and refolding of apomyoglobin in vacuum: An activation barrier for gas-phase protein folding. *J. Am. Chem. Soc.* **119**, 2987-2994, (1997).

337   Bohrer, B. C., Merenbloom, S. I., Koeniger, S. L., Hilderbrand, A. E. & Clemmer, D. E. Biomolecule analysis by ion mobility spectrometry. *Annu. Rev. Anal. Chem.* **1**, 293-327, (2008).

338   Rizzo, T. R., Stearns, J. & Boyarkin, O. V. Spectroscopic studies of cold, gas-phase biomolecular ions. *Int. Rev. Phys. Chem.* **28**, 481-515, (2009).

339   Scott, A. P. & Radom, L. Harmonic vibrational frequencies: An evaluation of hartree–fock, møller–plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *J. Phys. Chem.* **100**, 16502-16513, (1996).

340   Bouteiller, Y., Gillet, J.-C., Grégoire, G. & Schermann, J. P. Transferable specific scaling factors for interpretation of infrared spectra of biomolecules from density functional theory. *J. Phys. Chem. A* **112**, 11656-11660, (2008).

341   Vendrell, O., Gatti, F. & Meyer, H.-D. Full dimensional (15-dimensional) quantum-dynamical simulation of the protonated water dimer. II. Infrared spectrum and vibrational dynamics. *J. Chem. Phys.* **127**, 184303, (2007).

342   Rossi, M., Ceriotti, M. & Manolopoulos, D. E. How to remove the spurious resonances from ring polymer molecular dynamics. *J. Chem. Phys.* **140**, 234116, (2014).

343   Zabuga, A. V. & Rizzo, T. R. Capping motif for peptide helix formation. *The Journal of Physical Chemistry Letters* **6**, 1504-1508, (2015).

344   Bush, M. F., O'Brien, J. T., Prell, J. S., Saykally, R. J. & Williams, E. R. Infrared spectroscopy of cationized arginine in the gas phase : Direct evidence for the transition from nonzwitterionic to zwitterionic structure. *J. Am. Chem. Soc.* **129**, 1612-1622, (2007).

345   Polfer, N. C. & Oomens, J. Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance. *Mass Spectrom. Rev.* **28**, 468-494, (2009).

346   Chiavarino, B. *et al.* Cation-pi interactions in protonated phenylalkylamines. *J. Phys. Chem. A* **118**, 7130-7138, (2014).

347   Eliezer, D. in *Methods in molecular biology: Protein folding protocols* Vol. 350 (eds Y. Bai & R. Nussinov) 49-67 (Humana Press Inc., 2007).

348   Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* **5**, 67-81, (1995).

349   Montaldi, L. R., Berardi, M., Souza, E. S., Juliano, L. & Ito, A. S. End-to-end distance distribution in fluorescent derivatives of bradykinin in interaction with lipid vesicles. *J. Fluoresc.* **22**, 1151-1158, (2012).

350   Voronina, L. & Rizzo, T. R. Spectroscopic studies of kinetically trapped conformations in the gas phase: The case of triply protonated bradykinin. *Phys. Chem. Chem. Phys.* **17**, 25828-25836, (2015).

351   Kaplan, A. P., Joseph, K. & Silverberg, M. Pathways for bradykinin formation and inflammatory disease. *J. Allergy Clin. Immunol.* **109**, 195-209.

352   Rhaleb, N.-E. *et al.* Structure-activity studies of bradykinin and related peptides. *Hypertension* **17**, 107-115, (1997).

353   Marceau, F., Lussier, A., Regoli, D. & Giroud, J. P. Pharmacology of kinins: Their relevance to tissue injury and inflammation. *Gen. Pharmacol.* **14**, 209-229, (1983).

354   Sheikh, I. A. & Kaplan, A. P. Studies of the digestion of bradykin, Lys-bradykinin, and des-Arg9-bradykinin by angiotensin converting enzyme. *Biochem. Pharmacol.* **35**, 1951-1956, (1986).

355   Kemper, P. R., Dupuis, N. F. & Bowers, M. T. A new, higher resolution, ion mobility mass spectrometer. *Int. J. Mass spectrom.* **287**, 46-57, (2009).

356   London, R. E., Stewart, J. M., Cann, J. R. & Matwiyoff, N. A. 13C and proton nuclear magnetic resonance studies of bradykinin and selected peptide fragments. *Biochemistry* **17**, 2270-2277, (1978).

357   Pierson, N. A., Chen, L., Valentine, S. J., Russell, D. H. & Clemmer, D. E. Number of solution states of bradykinin from ion mobility and mass spectrometry measurements. *J. Am. Chem. Soc.* **133**, 13810-13813, (2011).

358   Shvartsburg, A. A. *et al.* Characterizing the structures and folding of free proteins using 2-D gas-phase separations: Observation of multiple unfolded conformers. *Anal. Chem.* **78**, 3304-3315, (2006).

359   Tang, K., Li, F., Shvartsburg, A. A., Strittmatter, E. F. & Smith, R. D. Two-dimensional gas-phase separations coupled to mass spectrometry for analysis of complex mixtures. *Anal. Chem.* **77**, 6381-6388, (2005).

360   Papadopoulos, G., Svendsen, A., Boyarkin, O. V. & Rizzo, T. R. Spectroscopy of mobility-selected biomolecular ions. *Faraday Discuss.* **150**, 243-243, (2011).

361   Nagornova, N. S., Rizzo, T. R. & Boyarkin, O. V. Highly resolved spectra of gas-phase gramicidin s: A benchmark for peptide structure calculations. *J. Am. Chem. Soc.* **132**, 4040-4041, (2010).

362   Stearns, J. A. *et al.* Conformation-specific spectroscopy and photodissociation of cold, protonated tyrosine and phenylalanine. *J. Am. Chem. Soc.* **129**, 11814-11820, (2007).

363 Doemer, M. *et al.* Assessing the performance of computational methods for the prediction of the ground state structure of a cyclic decapeptide. *Int. J. Quantum Chem* **113**, 808-814, (2013).

364 Rossi, M. *et al.* Secondary structure of Ac-Ala(n)-LysH(+) polyalanine peptides (n=5, 10, 15) in vacuo: Helical or not? *J. Phys. Chem. Lett.* **1**, 3465-3470, (2010).

365 Baldauf, C. *et al.* How cations change peptide structure. *Chem. - Eur. J.* **19**, 11224-11234, (2013).

366 Jaeqx, S., Oomens, J., Cimas, A., Gaigeot, M. P. & Rijs, A. M. Gas-phase peptide structures unraveled by far-IR spectroscopy: Combining IR-UV ion-dip experiments with Born-Oppenheimer molecular dynamics simulations. *Angew. Chem.* **53**, 3663-3666, (2014).

367 Stewart, D. E., Sarkar, A. & Wampler, J. E. Occurrence and role of cis peptide bonds in protein structures. *J. Mol. Biol.* **214**, 253-260, (1990).

368 Brazin, K. N., Mallis, R. J., Fulton, D. B. & Andreotti, A. H. Regulation of the tyrosine kinase Itk by the peptidyl-prolyl isomerase cyclophilin A. *Proc. Natl. Acad. Sci. USA* **99**, 1899-1904, (2002).

369 Royer, C. A. *et al.* Effects of amino acid substitutions on the pressure denaturation of staphylococcal nuclease as monitored by fluorescence and nuclear magnetic resonance spectroscopy. *Biochemistry* **32**, 5222-5232, (1993).

370 Williams, A. D. *et al.* Mapping a β amyloid fibril secondary structure using scanning proline mutagenesis. *J. Mol. Biol.* **335**, 833-842, (2004).

371 Hacke, M., Gruber, T., Schulenburg, C., Balbach, J. & Arnold, U. Consequences of proline-to-alanine substitutions for the stability and refolding of onconase. *FEBS J.* **280**, 4454-4462, (2013).

372 Kang, K. *et al.* The three proline residues (P25, P242, and P434) of agrobacterium CP4 5-enolpyruvylshikimate-3-phosphate synthase are crucial for the enzyme activity. *Plant Biotechnol. Rep.* **4**, 329-334, (2010).

373 Chao, H., Davies, P. L., Sykes, B. D. & Sönnichsen, F. D. Use of proline mutants to help solve the NMR solution structure of type III antifreeze protein. *Protein Sci.* **2**, 1411-1428, (1993).

374 Dumy, P. *et al.* Pseudo-prolines as a molecular hinge: Reversible induction of cis amide bonds into peptide backbones. *J. Am. Chem. Soc.* **119**, 918-925, (1997).

375   Keller, M. *et al.* Enhancing the proline effect: Pseudo-prolines for tailoring cis/trans isomerization. *J. Am. Chem. Soc.* **120**, 2714-2720, (1998).

376   Lorenzen, S., Peters, B., Goede, A., Preissner, R. & Frommel, C. Conservation of cis prolyl bonds in proteins during evolution. *Proteins* **58**, 589-595, (2005).

377   Raulfs, M. D. M. *et al.* Investigations of the mechanism of the "proline effect" in tandem mass spectrometry experiments: The "pipecolic acid effect". *J. Am. Soc. Mass. Spectrom.* **25**, 1705-1715, (2014).

378   Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95-99, (1963).

379   Kleywegt, G. J. & Jones, T. A. Phi/psi-chology: Ramachandran revisited. *Structure* **4**, 1395-1400, (1996).

380   Shi, Z., Chen, K., Liu, Z. & Kallenbach, N. R. Conformation of the backbone in unfolded proteins. *Chem. Rev.* **106**, 1877-1897, (2006).

381   Chakrabarti, P. & Pal, D. The interrelationships of side-chain and main-chain conformations in proteins. *Prog. Biophys. Mol. Biol.* **76**, 1-102, (2001).

382   Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. Procheck: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291, (1993).

383   Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* **381**, 272-272, (1996).

384   Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. Stereochemical quality of protein structure coordinates. *Proteins: Struct. Funct. Bioinform.* **12**, 345-364, (1992).

385   Anderson, R. J., Weng, Z., Campbell, R. K. & Jiang, X. Main-chain conformational tendencies of amino acids. *Proteins: Struct. Funct. Bioinform.* **60**, 679-689, (2005).

386   Lietz, C. B. *et al.* Multiple gas-phase conformations of proline-containing peptides: Is it always cis/trans isomerization? *Analyst* **141**, 4863-4869, (2016).

387   Schenk, E. R., Ridgeway, M. E., Park, M., Leng, F. & Fernandez-Lima, F. Isomerization kinetics of at hook decapeptide solution structures. *Anal. Chem.* **86**, 1210-1214, (2014).