Check for updates

Chemistry–A European Journal
Research Article
doi.org/10.1002/chem.202303715
Chemistry Europe
European Chemical Societies Publishing
www.chemeurj.org

Binding of Acetylated Lysine by Using a Water Soluble Aryl Extended Calix[4]pyrrole

Martina Orlandini,[a] Alessandro Pedrini,[a] Danilo Marchetti,[a, b] Yifan Li,[c, d] Gemma Aragay,[c] Enrico Dalcanale,[a] Pablo Ballester,*[c, e] and Roberta Pinalli*[a]

Post-translational modifications of lysine in histones, as methylation and acetylation, have well established functions in epigenetics and are emerging as important actors in broader biological regulation. Currently, the detection of acetylated lysine (**Kac**) in water solution as free amino acid or protein residue remains challenging. Acetylated lysine is a neutral amino acid, and the lack of ion–dipole interactions causes the decrease in binding affinity displayed by synthetic molecular receptors with respect to the other lysine modifications. Here, we report molecular modeling calculations and ^{1}H NMR experiments to investigate the binding properties of two different calix[4]pyrrole receptors towards **Kac**. Computational analyses reveal that tetra-aryl-extended calix[4]pyrrole (**1**) preferentially binds the *cis*-**Kac** conformer over the *trans* one due to steric considerations and more favorable interactions. Experimental ^{1}H NMR titration experiments confirm the formation of a 1:1 complex between receptor **1** and *cis*-**Kac**, with a K_a exceeding 10^3 M^{-1}. Conversely, the super-aryl-extended calix[4]pyrrole **2** is less efficient in binding **Kac**, due to unfavorable solvation/desolvation effects, as proven by ^{1}H NMR experiments. Moreover, receptor **1** showed a higher affinity for **Kac** over other lysine modifications, such as methylated lysines.

Introduction

Post-translational modifications (PTMs) of proteins, such as acetylation, methylation, phosphorylation and ubiquitylation, play essential roles in cell biology, including the regulation of gene expression and chromatin dynamics, protein-protein interactions, localization, and stability, as well as enzymatic activities of proteins involved in many cellular processes.[1] In particular, acetylation of lysine residues (Kac) has emerged as a crucial PTM for a wide range of cellular processes and it is involved in aging and in the development of several diseases, including leukemia.[2] Lysine acetylation is a dynamic and reversible process in which an acetyl group from acetyl CoA is covalently bonded to the ε-position of a lysine residue through enzymatic or non-enzymatic means.[3] Being a secondary amide, an acetylate lysine can exist as two different conformers, the *cis* and the *trans* amide bond rotamers (Scheme 1b), depending on the relative position of the carbonyl group with respect to the amide N–H. For acetylated lysines, the calculations assigned a stabilization energy of 2.4–2.5 kcal/mol to the *trans*-rotamer with respect to the *cis* counterpart. A similar tendency of energies has been reported in the literature for other secondary amides.[4,5,6,7] The free energy barrier for the *cis-trans* interconversion is known to be moderate ($\Delta G = 16–22$ kcal mol^{-1}).[8] Therefore, the process is fast on the human time-scale but slow enough on the ^{1}H NMR chemical shift timescale to allow the detection of separate signals for the two rotamers.[6]

Mass spectrometry is the most used technique to identify type and site of protein modifications.[9] The analysis of PTMs represents a challenge since most of them are present in low stoichiometry, few residues in the target protein contain the modification and some PTMs produce signal suppression and low-quality MS/MS spectra.[10] Therefore, there is an urgent need for the development of efficient enrichment and purification methods to improve the identification of modified peptides containing acetylated lysine residues.

One of the most common methods used to enrich Kac peptides relies on an antibody-based immunoaffinity enrichment.[11] The method is effective in Kac enrichment, however, the use of antibodies is expensive and suffers from batch-to-batch reproducibility.[12] A novel approach to improve the enrichment of Kac peptides in real samples was described by Zhang and co-workers.[13] The authors proposed the use of a

[a] Dr. M. Orlandini, Dr. A. Pedrini, D. Marchetti, Prof. E. Dalcanale, Prof. R. Pinalli
Department of Chemistry, Life Sciences and Environmental Sustainability
University of Parma
Parco Area delle Scienze 17/A, 43124 Parma, Italy
E-mail: roberta.pinalli@unipr.it

[b] D. Marchetti
Istituto Italiano di Tecnologia, Center for Materials Interfaces, Electron Crystallography
Viale Rinaldo Piaggio 34, 56025 Pontedera, Italy

[c] Y. Li, Dr. G. Aragay, Prof. P. Ballester
Institute of Chemical Research of Catalonia (ICIQ), Barcelona Institute of Science and Technology (BIST)
Avgda. Països Catalans 16, 43007 Tarragona, Spain
E-mail: pballester@iciq.es

[d] Y. Li
Departament de Química Analítica i Química Orgànica
Universitat Rovira i Virgili (URV)
c/Marcel·lí Domingo 1, 43007 Tarragona, Spain

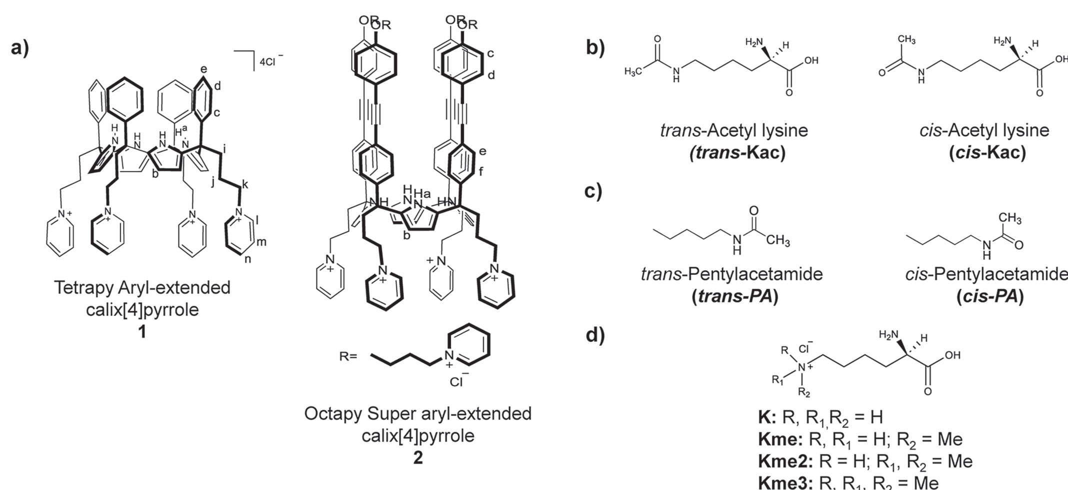
[e] Prof. P. Ballester
ICREA
Pg. Lluís Companys, 23, 08018 Barcelona, Spain

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202303715

© 2023 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

Chemistry Europe
European Chemical Societies Publishing



Scheme 1. Molecular structures of **a)** Tetrapyridinium aryl-extended calix[4]pyrrole receptor **1** and octapyridinium super aryl-extended calix[4]pyrrole receptor **2**; **b)** *trans*- and *cis*- conformers of acetyl lysine (*trans*-**Kac** and *cis*-**Kac**); **c)** *trans*- and *cis*-conformers of pentylacetamide (*trans*-**PA** and *cis*-**PA**); and **d)** lysine, mono-, di and tri-methyl lysine (**K**, **Kme**, **Kme2** and **Kme3**, respectively).

molecularly imprinted monolith as a tool for the identification of Kac-modification and peptide fractionation. The imprinted material employed a well-defined peptide sequence as template to efficiently capture the corresponding acetylated peptides. Therefore, distinct peptide sequences containing the Kac residue are needed as templates to prepare a useful portfolio of molecular imprinted materials capable of an efficient enrichment of the corresponding peptides.

Supramolecular receptors represent a valid alternative to the use of antibodies or molecular imprinted materials for the recognition and binding of specific guests. Macrocycles like cyclodextrins, calixpyrroles, calixarenes, cucurbiturils and cavitands are prominent representatives of synthetic receptors, featuring excellent molecular recognition properties for a large and assorted pool of guests. The molecular recognition properties of the synthetic receptors are generally well-studied both in terms of analytes of interest and interferents. In order to improve macrocycles selectivity and binding affinity, the size and shape of their cavities, as well as the type, number and directionality of the interactions can be tuned. Likewise, the principles of preorganization and complementarity for an efficient molecular recognition process must be taken into account.[12] The molecular recognition properties of synthetic receptors have already been exploited for the targeting of simple amino acids, as well as protein modifications.[12] For example, *p*-sulfonatocalix[4]arene strongly binds trimethyl lysine with good selectivity both as a free amino acid and as an exposed residue in a peptide chain.[14] Cucurbit[7]uril efficiently binds tri-, di- and mono-methylated lysine,[15] as well as a dimethylated lysine residue in a model protein.[16] In 2016, Pinalli et al. explored the ability of tetraphosphonate cavitands to selectively bind mono-methylated lysine in water.[17] Later on, this ability of the tetraphosphonate receptor was transferred to the detection of lysine mono-methylation in human histone tails.[18]

To date, the detection of acetylated lysine in water solution as free amino acid or as a protein residue remains challenging. Acetylated lysine is a neutral amino acid, and the lack of ion–dipole interactions account for the decrease in binding affinity displayed by molecular receptors like calixarenes[14] or cucurbiturils[15] with respect to the other lysine modifications providing charged-compounds. In 2018, Ballester and co-workers proposed the use of a water soluble tetra-α aryl-extended calix[4]pyrrole as molecular receptor for the binding of primary acetamides with high affinity. The amides are bound in the endo-functionalized aromatic cavity of the calix[4]pyrrole by a combination of the hydrophobic effect, hydrogen-bonding, NH–π and CH–π interactions. Remarkably, the receptor displayed a high conformational selectivity for the amide *cis* rotamer.

Based on the above findings, we decided to explore the ability of a calix[4]pyrrole for the binding of acetylated lysine in aqueous solution. Herein, we report the results of a fundamental study on the molecular recognition properties of a tetra phenyl-extended calix[4]pyrrole toward acetylated lysine. We also disclose the biding selectivity of the receptor for potential interferents like unmodified and methylated lysines. The ability of the calix[4]pyrrole to bind Kac was investigated by molecular modelling studies and experimentally confirmed using ^{1}H NMR spectroscopy in water solution. This study validates the tetra-α isomer of *meso*-tetraphenyl calix[4]pyrrole **1** (Scheme 1a)) as a potential synthetic receptor for the selective recognition of acetylated lysine residues in proteins.

Results and Discussion

In this study, we used two different water-soluble calix[4]pyrroles: the tetra-α isomer of the tetra-pyridinium tetraphenyl calix[4]pyrrole **1**, bearing four *meso*-phenyl substituents, and the octa-pyridinium super aryl-extended calix[4]pyrrole **2**,

© 2023 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202303715 by Consorci De Serveis Universitaris De Catalunya, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Research Article
doi.org/10.1002/chem.202303715

Chemistry Europe
European Chemical Societies Publishing

presenting a deeper aromatic cavity defined by 4-(phenylethenyl)-phenyl *meso*-substituents (Scheme 1a)). Water solubility is guaranteed by the presence of four pyridinium groups at the lower rim for receptor **1**, and eight pyridinium groups, four at the upper and four at the lower rims, for receptor **2**. Calix[4]pyrroles **1** and **2** were prepared following literature procedures (see SI for details).[8,19,20,21]

The molecular structures of the guests employed in this study are depicted in Scheme 1b)—d). Acetylated lysine (**Kac**) is the target guest, while pristine lysine and mono-, di- and trimethylated lysines (**K**, **Kme**, **Kme2** and **Kme3**, respectively) are used as competitive guests to evaluate the selectivity of the receptors toward the former. Finally, we used pentyl acetamide (**PA**) to dissect the influence of steric and solvation/desolvation effects in the binding of **Kac**.

Molecular modelling studies

As already discussed, molecular receptors based on a tetra-α aryl-extended calix[4]pyrrole scaffold are known to exhibit a selective binding toward the *cis*-rotamers of acetamides.[8] The ability of receptors **1** and **2** to complex acetylated lysine was firstly analyzed via computational studies (See SI, Section 2). For receptor **1**, molecular modelling calculations were performed on the inclusion complexes of both *cis* and *trans*-**Kac**, considering that the guest was included in the aromatic cavity by establishing four convergent hydrogen bonds between the oxygen atom and the pyrrole NHs of the host (Figure 1). The energies of the host-guest complexes revealed that the binding of the *cis*-**Kac** conformer was favored by 9.1 kcal·mol^{-1}. At first glance, the analysis of the energy minimized structure of the *trans*-**Kac@1** complex (Figure 1b) showed the existence of steric clashes between the amino-acid alkyl chain and the aromatic walls of the calix[4]pyrrole receptor in cone conformation. This produces a widening in the open end of the receptor **1** cavity (Figure S5).

Two of the axially oriented *meso*-phenyl substituents of the host bend toward the exterior of the aromatic cavity to diminish the steric clashes. The resulting conformation of the *trans*-**Kac** complex resembles that of a resorcin[4]arene pinched-cone. Conversely, the *cis*-**Kac@1** complex showed a better fit of the guest's dimensions to those of the receptor cavity in an ideal cone conformation (Figure 1a).

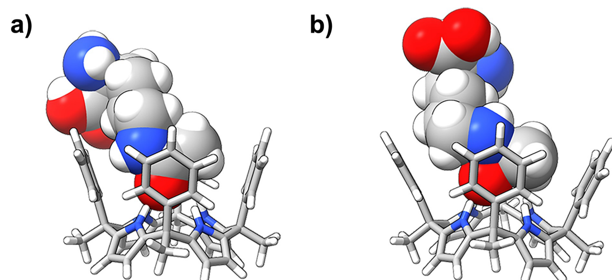


Figure 1. Energy minimized models of the receptor **1** and **Kac** complexes: **a)** *cis*-**Kac@1**; **b)** *trans*-**Kac@1**. Geometrical optimization was performed using a semi-empirical PM3 level of theory, while the energy calculation utilized a DFT-B3LYP/6-31+(d) level of theory. The guest molecules are shown as CPK models, receptor **1** in stick representation and the *meso*-(3-pyridinium-propyl) groups have been omitted for clarity.

As mentioned above, in both complexes the **Kac** establishes four hydrogen bond interactions between its carbonyl oxygen atom and the pyrrole NHs of the host. Moreover, the two inclusion complexes are stabilized by additional N–H···π and CH–π interactions established between the amide substituents and the aromatic panels of **1**[22,23,24] (Figure S4). The computational results related to receptor **2** are reported in SI.

^{1}H NMR spectroscopy experiments

Experimentally, the molecular recognition properties of receptors **1** and **2** toward **Kac** were evaluated using ^{1}H NMR titration experiments, adding incremental amounts of **Kac** [40 mM] into a D_2O solution of the receptor [2.6 mM]. **Kac** is present in solution as a mixture of two rotamers in equilibrium (*cis*-**Kac** and *trans*-**Kac**, Scheme 1b)). As already discussed in the introduction section, this equilibrium features slow chemical exchange dynamics on the chemical shift timescale. Therefore, the ^{1}H NMR spectrum of free **Kac** shows separate singlets of different intensity for the methyl protons of the acetyl group in their two different rotamers (Figures 2a and S10). According to the literature,[4,5,7,25] we assigned the lower intensity singlet resonating at δ=2.06 ppm to the *cis* conformer of **Kac** (H_1) and the more intense singlet appearing at δ=1.99 ppm was associated to the *trans* rotamer of **Kac** (H_2). The proton assignment was also supported by calculated chemical shift values for the two conformers using Gaussian (DFT theory level B3LYP/6-311+(3d,p)). The computed values were δ=2.1 and

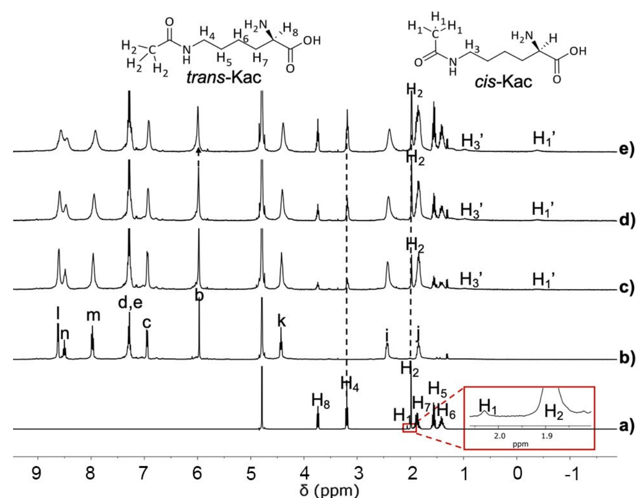


Figure 2. Top) Molecular structure of *cis*- and *trans*- **Kac** with the corresponding proton assignment. Bottom) ^{1}H NMR (400 MHz, D_2O, 298 K) spectra acquired during the titration of a D_2O solution of **1** [2.6 mM] with incremental amounts of a D_2O solution of **Kac** [40 mM]. **a)** **Kac**, **b)** receptor **1**, **c)** **1**+1 eq of **Kac**, **d)** **1**+2 eq of **Kac**, **e)** **1**+5 eq of **Kac**. H_1' and H_3' protons are referred to bound *cis*-**Kac**. Proton signals for the chain of the free *cis*-**Kac** have similar chemical shift values to the *trans*-**Kac**, but they are not detectable due to the low concentration of the cis rotamer. The concentration of **1** was kept constant during the titration. The full ^{1}H NMR titration experiment is reported in Figure S11.

© 2023 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202303715 by Consorci De Serveis Universitaris De Catalunya, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Research Article
doi.org/10.1002/chem.202303715

1.9 ppm for the methyl protons of the acetyl groups in *cis*-**Kac** and *trans*-**Kac** rotamers, respectively (Table S2). The *cis/trans* ratio of the rotamers was experimentally calculated from the integral values of their methyl signals, and it resulted to be ~1/99 (*cis*-**Kac**/*trans*-**Kac** ratio) (Figure S10), which is in line with the values reported in literature.[7]

^{1}H NMR complexation experiment using receptor 1 as host and Kac as guest

Firstly, we performed the titration of receptor **1** with **Kac** in neutral D_2O (Figure 2). The concentration of receptor **1** was maintained constant throughout the titration.

The addition of incremental amounts of **Kac** to a 2.6 mM solution of **1**, produced upfield shifts to the proton signals assigned to H_l, H_n, H_m, H_d, H_e and H_c in the free receptor **1** (see Scheme 1 for proton assignment). In contrast, the singlet H_b assigned to the β-pyrrole protons moved downfield. Most of the proton signals of the receptor **1** broadened after the addition of the guest. This observation suggested that the equilibrium binding process featured intermediate chemical exchange dynamics on the chemical shift timescale. Throughout the titration, the singlet of the methyl group of the *trans*-**Kac** rotamer remained sharp and did not show a noticeable chemical shift change (H_2, Figure 2). This observation suggests that *trans*-**Kac** is not bound to receptor **1**. Remarkably, even after the addition of an excess of **Kac** (5 eq.), we could not detect the proton signal assigned to the methyl group of the free *cis*-**Kac** (Figure 2e). Instead, we observed the appearance of new broad signals at $\delta = -0.48$ ppm and at $\delta = 0.97$ ppm that grew in intensity upon increasing the concentration of the guest (H_1' and H_3', Figure 2c–e). We assigned the peak at $\delta = -0.48$ ppm to the methyl group of the *cis*-**Kac** isomer (H_1') bound in the polar aromatic cavity of **1**, while the peak at $\delta = 0.97$ ppm to the CH_2 in alpha position to the amino group of the amide (H_3'). The singlet corresponding to free *cis*-**Kac** rotamer (H_1) broadened beyond detection (H_1') due intermediate exchange on the chemical shift timescale with the bound counterpart. The chemical exchange between the two free rotamers remains slow on the chemical shift time scale. Most likely, the rest of the peaks of the *cis*-**Kac** chain lies below the signals of both receptor **1** and *trans*-**Kac** resonating in the aliphatic region of the spectrum. The complexation induced shift (CIS) experienced by the methyl group of the *cis*-rotamer was consistent with that reported in literature for *cis*-rotamers of acetamides included in receptor **1**.[8] Moreover, the computed chemical shift value for the methyl group in the *cis*-**Kac@1** complex (DFT theory level B3LYP/6-311+(3d,p), Table S2) also supported this conclusion. Unfortunately, we did not detect cross peaks due to chemical exchange between the CH_3 signals of the acetyl group for the free and bound *cis* rotamer in the 2D-NMR experiments (NOESY) performed at the end the titration (5 eq. of **Kac**). Either the broadening and reduced intensity of the signals or the existence of a chemical exchange that was slow on the 2D experiments timescale hampered the observation of chemical exchange cross-peaks. The chemical shift changes experienced by the proton H_b of receptor **1** were fit to a theoretical binding model considering the existence of an equilibrium between the two rotamers in solution ($K_{eq} = 1\times10^{-2}$, for *cis*-**Kac**/*trans*-**Kac** = 1/99), and the exclusive formation of a 1:1 complex between **1** and the *cis*-**Kac** conformer. We determined a binding constant for the *cis*-**Kac@1** complex larger than 10^3 M^{-1} (see additional Supporting Information document). This value was in line with similar host-guest complexes reported in the literature.[8] We also estimated a binding constant for the putative 1:1 complex using integral values of the signals attributed to the bound *cis*-**Kac** and those of receptor **1**. The estimated value for the *cis*-**Kac@1** complex (6609 M^{-1}, see SI, Section 3 for calculation) is of the same order of magnitude than the one calculated from the fit of the chemical shift changes experienced by the β-pyrrole protons (see SI and Figure S12). Not surprisingly, Isothermal Titration Calorimetry experiments were not successful for determining an accurate binding constant (see SI, Section 4 for details).

^{1}H NMR complexation experiment using Kac as guest and receptor 2 as host

Next, we studied the binding of **Kac** with receptor **2**, featuring a much deeper hydrophobic polar cavity respect to receptor **1** (Scheme 1a) and Figure S15). The ^{1}H NMR spectrum of **2** in D_2O showed broad and ill-defined proton signals owing to the aggregation tendency of the receptor in water (Figure S15b). We expected that the formation of the *cis*-**Kac@2** complex would reduce the aggregation tendency of the bound receptor, resulting in a sharpening of its proton signals. However, the addition of increasing amounts of **Kac** to a 1 mM solution of **2** in D_2O did not produce significant changes in the proton signals of both host and guest. This result suggested that receptor **2** is not a suitable host for the **Kac** rotamers.

We speculated that the binding preference of the *cis*-**Kac** rotamer for receptor **1** over receptor **2** could be related to the solvation/desolvation effects experienced by the amino-acid polar head groups when included in **2**. In the *cis*-**Kac@1** complex, the amino-acid head groups are exposed to the bulk solvent (Figure 1a)). In addition, the lack of size and functional complementarity between the polar head groups and the cavity defined by **2**, suggested a reduced thermodynamic stabilization for the *cis*-**Kac@2** complex.

As highlighted in the energy-minimized structure of complex *cis*-**Kac@1** (Figure 1a)), the hydrophobic cavity of receptor **1** is well suited to accommodate the lysine aliphatic chain establishing four hydrogen bond interactions between the oxygen atom of the acetyl C=O and the four NHs of the pyrrole units (Figure S4). The aminoacidic part lies outside the cavity and it is still exposed to the water solvation. On the other hand, the energy minimized structure of *cis*-**Kac@2** complex (Figure S9) shows that the complexation event will require larger desolvation of the included guest, which enters completely inside the receptor cavity.

To support our hypothesis, we performed a ^{1}H NMR spectroscopic titration experiment of receptor **2** with pentylace-

© 2023 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202303715 by Consorci De Serveis Universitaris De Catalunya, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202303715 by Consorci De Serveis Universitaris De Catalunya, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

tamide (**PA**, Scheme 1c), the synthesis is reported in the SI). As **Kac**, **PA** is a secondary acetamide and exists in solution as two isomeric *cis-trans* rotamers. It presents an aliphatic chain of similar length than that of lysine but lacks the terminal carboxylic and amino head groups. The formation of the putative *cis*-**PA@2** complex should not involve energetic penalties related to solvation/desolvation of the polar head groups of the guest.

^{1}H NMR control experiment using PA as guest and receptor 2 as host

We performed a ^{1}H NMR titration experiment in D_2O solution, by adding incremental amounts of **PA** [45 mM] to a millimolar D_2O solution of receptor **2** [1.4 mM] (Figure 3).

The ^{1}H NMR spectrum of **PA** (Figure S16 and 3a) showed a *cis/trans* ratio of rotamers of ~1/99, calculated from the integral values of their methyl peaks (Fugure S16, grey and yellow dots for the *cis* and the *trans* isomer, respectively).[24]

As discussed above, the ^{1}H NMR spectrum of **2** displayed very broad signals due to aggregation (Figure 3b). The addition of increasing amounts of **PA** induced the sharpening of receptors signals (Figures 3c–e). Moreover, (Figure 3c–e), a new set of growing proton signals was observed in the upfield region (− 0.6 ppm to 0.5 ppm) of the ^{1}H NMR spectra acquired during the titration. We assigned this new set of signals to the protons of the aliphatic chain and the methyl protons of the *cis*-**PA** included in the cavity of **2** (H_{10}', H_9', H_8' and H_1', respectively). As expected, the proton signals of the *trans*-**PA** isomer did not experience noticeable changes (H_3, H_4, H_5, H_6, H_7), and their intensity grew with the increase of the guest concentration.

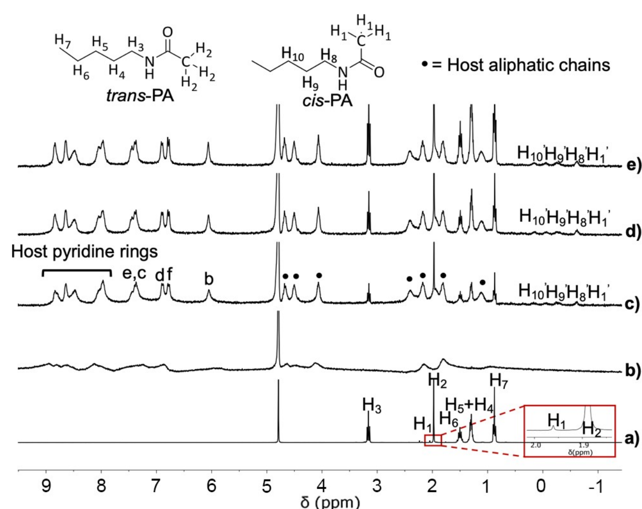


Figure 3. ^{1}H NMR (400 MHz, D_2O, 298 K) spectra acquired during the titration of incremental amounts of **PA** [45 mM] into a D_2O solution of **2** [1.4 mM]. **a)** **PA**, **b)** receptor **2**, **c)** 2+0.5 eq of **PA**, **d)** 2+1 eq of **PA**, **e)** 2+2 eq of **PA**. Primed letters indicate protons signals related to bound *cis*-**PA**. Proton signals for the chain of the free *cis*-**PA** have similar chemical shift values to the *trans*-**PA**, but they are not detectable due to the low concentration of the cis rotamer. The concentration of **2** was kept constant during the titration. The full ^{1}H NMR titration experiment is reported in Figure S**17**.

Throughout the titration, we did not detect the signals corresponding to free rotamer of *cis*-**PA** (Figure S17). Most likely, this was due to the broadening of the signals caused by the chemical exchange between free and bound *cis*-**PA**, as well as the overlap of the NH signal of free *cis*-**PA** with those of **2**.

The *cis*-**Kac@1** complex exclusively displayed significant upfield shifts for the protons nearby the amide bond (e.g. acetyl CH_3 and CH_2 next to the NH) because only this part of the aliphatic chain was included within the cavity of **1** and experienced the shielding effect of the *meso*-phenyl groups. In contrast, for the *cis*-**PA@2** complex, owing to the deeper cavity defined by the 4-(phenylethenyl)-phenyl *meso*-substituents, the methylene protons of the alkyl chain of bound *cis*-**PA** are also included in the aromatic cavity of **2** experiencing related shielding effects and significan upfield shifts. The addition of more than 1 eq. of **PA** did not produce significant changes in the shape and chemical shift values of the proton signals of receptor **2**. Therefore, we estimated that the binding constant of the *cis*-**PA@2** complex was larger than 10^4 M^{-1}.

The results of the binding of *cis*-**PA** with **2** demonstrated the deleterious effect caused by the amino acid head groups in the analogous binding of *cis*-**Kac**. Most likely, hots-guest shape complementarity and the establishment of suitable interactions were not sufficient for the formation of a thermodynamically stable supramolecular complex *cis*-**Kac@2**. In addition, not favourable solvation/desolvation effects were also playing a crucial role in the binding event of *cis*-**Kac** by **2**.

Selectivity studies

Once we proved that receptor **1** could effectively bind *cis*-**Kac** in water, we studied its binding selectivity for **Kac** over unmodified lysine (**K**, Scheme 1d)) and other epigenetic PTMs biomarkers like mono-, di- and tri-methyl lysine (**Kme**, **Kme2** and **Kme3**, respectively, Scheme 1d)) as possible competitive guests.

Both molecular modelling studies and ^{1}H NMR experiments were used to assess the ability of receptor **1** in binding the guests' series.

Molecular modelling

Molecular modelling calculations (DFT theory level B3LYP/6-31 +(d)) on the series of methylated lysines highlighted the impossibility of optimizing the guest position inside the host cavity. This lack of affinity is related to the absence of specific interaction and the presence of electrostatic repulsion forces (Figure S3). In addition, the complexation of **Kme3** is also hampered by a wider steric hindrance with respect to the other guest molecules. We further investigated the complexes' energies considering the guests in their neutral form. According to the molecular modelling results, in their neutral form **K**, **Kme** and **Kme2** are feasible candidates for insertion into the receptor **1** aromatic cavity (Figure S7). Although the calculation led to acceptable complex energies for **K@1**, **Kme@1** and

© 2023 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202303715 by Consorci De Serveis Universitaris De Catalunya, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Kme2@1 (Table S3) as neutral guests, they established weaker interactions than the *cis*-acetyl lysine molecule (Table S1).

^{1}H NMR spectroscopy titrations

To confirm the results obtained from the molecular modelling studies, we performed ^{1}H NMR titration experiments of receptor **1** with pristine lysine and the methylated lysine derivatives. The addition of incremental amounts of a D_2O solution of **K** [45 mM] to a D_2O millimolar solution of receptor **1** [2.9 mM] did not produce significant changes to the proton signals of both **1** and **K** (Figure S18), even after the addition of an excess of guest. We observed an identical behavior in analogous titration experiments using **Kme** and **Kme2** as guests (Figures S19 and S20, respectively), indicating that **1** has no affinity for these guests. Considering the obtained results, we decided not to test the binding of **Kme3**.

Competitive ^{1}H NMR spectroscopy experiments using Kac and Kme as guest

To evaluate the hypothetical interference of PTMs biomarkers in the binding of *cis*-**Kac** by the receptor **1**, we performed a ^{1}H NMR pair-wise competitive experiment using equimolar amounts of both **Kac** and **Kme** in presence of receptor **1** as host (Figure 4).

During the experiment, we added 1 eq of **Kac** [44 mM in D_2O] to an equimolar solution of **1** and **Kme** (Figure 4c) in D_2O at millimolar concentration. As expected, the addition of **Kac** provoked the appearance of a broad signal at $\delta = -0.38$ ppm (Figure 4d). We assigned this signal to the protons of the methyl group of *cis*-**Kac** included in the cavity of **1**. The proton signals of **Kme** and the *trans*-**Kac** rotamer did not experience noticeable chemical shift change.

We performed the same experiment using **Kme2** as competitive guest obtaining identical results (Figure S21).

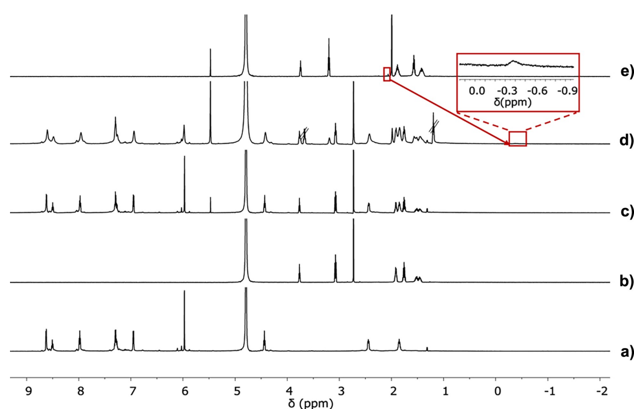


Figure 4. ^{1}H NMR (600 MHz, D_2O, 298 K) spectra acquired during the addition of **Kac** into a D_2O solution of **1** and **Kme**. a) receptor **1** [2.7 mM], b) **Kme** [42 mM], c) **1** + 1 eq of **Kme** d) **1** + 1 eq of **Kme** + 1 eq **Kac**, e) **Kac** [44 mM]. The concentration of **1** was kept constant during the titration.

Conclusions

We studied the molecular recognition properties of two different calix[4]pyrrole receptors, **1** and **2**, toward acetylated lysine via molecular modelling calculations and, experimentally, using ^{1}H NMR spectroscopic titrations in water. The tetra aryl-extended calix[4]pyrrole **1**, and the super aryl-extended calix[4]pyrrole **2**, featured four and eight pyridinium groups, respectively, as water solubilizing groups. In cone conformation, both receptors present a hydrophobic aromatic cavity with a polar endohedral binding site defined by four converging pyrrole NHs. Receptor **2** displays a deeper aromatic cavity compared to receptor **1**. We focussed our attention on the binding of acetylated lysine by the receptors since the former represents an important target residue in epigenetic field. As post translational modification, acetylated lysine is responsible of numerous human diseases.

Computational studies suggested that receptor **1** is suitable to complex acetylated lysine, and the calculated energies showed that the *cis*-**Kac@1** complex was energetically favored over the *trans*-**Kac@1** counterpart. Indeed, the energy minimized structures of the complexes showed a marked distortion of the host cavity in the *trans*-**Kac@1** complex. Probably, this was due to the larger steric hindrance of the amino-acid chain in the latter complex. Moreover, the *cis*-**Kac@1** complex was further stabilized by more favorable NH···O hydrogen bonds and additional NH···π interactions.

The results of ^{1}H NMR titrations demonstrated the formation of a 1:1 complex between **1** and the *cis*-**Kac** rotamer, while the *trans* rotamer did not bind the receptor. The association constant of the *cis*-**Kac@1** complex was derived from the chemical shift changes experienced by the proton of the host in the ^{1}H NMR spectra acquired for the titration experiment. The experimental data were fitted to a binding model considering the *cis/trans* isomerization equilibrium and the exclusive formation of a 1:1 complex between **1** and *cis*-**Kac** rotamer. We estimated that the K_a value of the *cis*-**Kac@1** complex was in the order of 10^3 M^{-1}.

Super-aryl extended calix[4]pyrrole **2** was not efficient for the binding of **Kac**. We hypothesized that the low binding affinity of **2** for **Kac** was mainly due to the energetically not favourable desolvation of the aminoacidic group that is required for its inclusion in the receptor's hydrophobic cavity. That is, the gain in entropy obtained from desolvation of the polar groups was not compensated by its enthalpic cost. To support our hypothesis, we performed ^{1}H NMR spectroscopic titration experiments using pentylacetamide (**PA**) as guest and receptor **2** as host. **PA** has a similar structure compared to **Kac**, but it lacks the amino-acidic terminal groups. Receptor **2** efficiently complexed the *cis*-**PA** rotamer. We estimated that the binding constant of the *cis*-**PA@2** complex was larger than 10^4 M^{-1}. This result confirmed that not only the presence of suitable interactions and host-guest shape complementarity are necessary for the formation of stable complexes, but also a suitable balance of the solvation/desolvation effects are fundamental.

The results of molecular modelling studies and ^{1}H NMR experiments revealed that other PTMs biomarkers as methylated

© 2023 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202303715 by Consorci De Serveis Universitaris De Catalunya, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

lysines were not bound by receptor **1**. The existence of electrostatic repulsion forces, and, in the case of **Kme3**, also steric hindrance are possible destabilizing factors of the complexes.

Tetra aryl-extended calix[4]pyrrole **1** resulted to be highly selective for the complexing of the *cis*-**Kac** rotamer of the free amino acid in the presence of lysine and its PTM derivatives (methylated lysine). The reported findings encourage the potential application of aryl-extended calix[4]pyrroles for the binding of acetylated lysine residues in peptides and proteins. Studies in this direction are currently on-going in our laboratories.

Supporting Information

The authors have cited additional references within the Supporting Information (Refs. [26–30]).

Acknowledgements

This work has benefited from the equipment and framework of the COMP-HUB and COMP-R Initiatives, funded by the 'Departments of Excellence' program of the Italian Ministry for University and Research (MIUR, 2018–2022 and MUR, 2023–2027). M. O. and R. P. thank University of Parma intramural funding "Bando di Ateneo per la Ricerca 2021". M. O, D. M., A. P., E. D. and R. P. acknowledge the Centro Interfacoltà di Misure "G. Casnati" of the University of Parma for the use of NMR facilities. Part of this research was funded by Ministerio de Ciencia e Innovación / Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033 (PID2020-114020GB-I00and-CEX2019-000925-S), the CERCA Programme / Generalitat de Catalunya, and AGAUR (2021SGR00851).

R. P. declares that this study was carried out within the «Innovative amplification strategies for ultrasensitive detection of molecular targets (P20200ANCEK, E53D23015640001) » project – funded by European Union – Next Generation EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This manuscript reflects only the authors' views and opinions and the Ministry cannot be considered responsible for them"

Conflict of Interests

The authors declare no conflict of interests.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: tetra-aryl-extended calix[4]pyrrole · acetyl lysine · molecular recognition · post translational modifications · solvation/desolvation effects

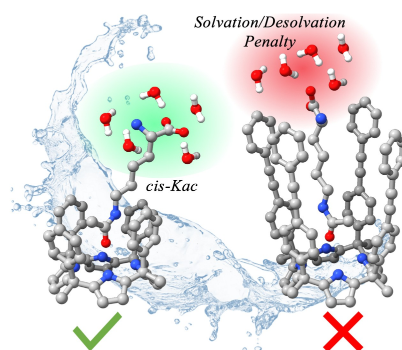
[1] a) X. Zhang, H. Wen, X. Shi, *Acta Biochem Biophys Sin* **2012**, *44*, 14–27; b) D. Levy, *Cell. Mol. Life Sci.* **2019**, *76*, 2873–2883.
[2] O. Van Beekum, E. Kalkhoven, *Subcell. Biochem.* **2007**, *41*, 233–262.
[3] I. Ali, R. J. Conrad, E. Verdin, M. Ott, *Chem. Rev.* **2018**, *118*, 1216–1252.
[4] G. Fischer, *Chem. Soc. Rev.* **2000**, *29*, 119–127.
[5] C. Dugave, L. Demange, *Chem. Rev.* **2003**, *10*, 2475–2532.
[6] W. E. Steward, T. H. Siddall, *Chem. Rev.* **1970**, *70*, 517–551.
[7] A. Genshaft, J. S. Moser, E. L. D'Antonio, C. M. Bowman, D. W. Christianson, *Proteins* **2013**, *81* 1051–1057.
[8] L. Escobar, A. Diaz-Moscoso, P. Ballester, *Chem. Sci.* **2018**, *9*, 7186–7192.
[9] a) M. I. Maron, S. M. Lehman, S. Gayatry, J. D. DeAngelo, S. Hegde, B. M. Lorton, Y. Sun, D. L. Bai, S. Sidoli, V. Gupta, M. R. Marunde, J. R. Bone, Z. W. Sun, M. T. Bedford, J. Shabanowitz, H. Chen, D. F. Hunt, D. Shechter, *iScience.* **2021**, *24*, 102971; b) S. Stransky, R. Cutler, J. Aguilan, E. Nieves, S. Sidoli, *Epigenetics Chromatin.* **2022**, *15*, 35.
[10] J. Gil, A. Ramìrez-Torres, S. Encarnacìon-Guevara, *Journal of Proteomics* **2017**, *150*, 297–309.
[11] T. Svinkina, H. Gu, J. C. Silva, P. Mertins, J. Qiao, S. Fereshetian, J. D. Jaffe, E. Kuhn, N. D. Udeshi, S. A. Carr, *Mol. Cell. Proteomics* **2015**, *14*, 2429–2440.
[12] R. Pinalli, A. Pedrini, E. Dalcanale, *Chem. Soc. Rev.*. **2018**, *47*, 7006.
[13] X. Zhang, Y. Wang, Z. H. Wei, D. Y. An, W. R. Pu, Z. S. Liu, Y. P. Huang, *Proteome Res.* **2022**, *21*, 325–338.
[14] C. S. Bashara, C. E. Jones, K. D. Daze, B. J. Lilgert, F. Hof, *ChemBioChem* **2010**, *11*, 63–66.
[15] M. A. Gamal-Eldina, D. H. Macartney, *Org. Biomol. Chem.* **2013**, *11*, 488–495.
[16] F. Guagnini, P. M. Antonik, M. L. Rennie, P. O'Byrne, A. R. Khan, R. Pinalli, E. Dalcanale, P. B. Crowley, *Angew. Chem. Int. Ed.* **2018**, *57*, 7126–7130.
[17] R. Pinalli, G. Brancatelli, A. Pedrini, D. Menozzi, D. Hernàndez, P. Ballester, S. Geremia, E. Dalcanale, *J. Am. Chem. Soc.* **2016**, *138*, 8569–8580.
[18] N. Bontempi, E. Biavardi, D. Bordiga, G. Candiani, I. Alessandri, P. Bergese, E. Dalcanale, *Nanoscale.* **2017**, *9*, 8639–8646.
[19] L. Escobar, G. Aragay, P. Ballester, *Chem. A Eur. J.* **2016**, *22*, 13682–13689.
[20] A. Dìaz-Moscoso, D. Hernandez-Alonso, L. Escobar, F. A. Arroyave, P. Ballester, *Org. Lett.* **2017**, *19*, 226–229.
[21] L. Escobar, P. Ballester, *Org. Chem. Front.* **2019**, *6*, 1738–1748.
[22] F. H. Beijer, R. P. Sijbesma, J. A. J. M. Vekemans, E. W. Meijer, H. Kooijman, A. L. Spek, *J. Org. Chem.* **1996**, *61*, 6371–6380.
[23] H. Adams, F. J. Carver, C. A. Hunter, N. J. Osborne, *Chem. Commun.* **1996**, 2529–2530.
[24] L. Alfonso, M. I. Burguete, F. Galindo, S. V. Luis, L. Vigara, *J. Org. Chem.* **2007**, *72*, 7947–7956.
[25] A. Radzicka, L. Pederson, R. Wolfenden, *Biochemistry* **1988**, *27*, 4538.
[26] D. B. G. Williams, M. Lawton, *J. Org. Chem.* **2010**, *75*, 8351–8354.
[27] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian, Inc., Wallingford CT, **2016**.
[28] J. B. Maglic, R. Lavendomme, *J. Appl. Crystallogr.* **2022**, *55*, 1033–1044.
[29] M. Inoue, C. Virués, I. Higuera-Ciapara, *Supramol. Chem.* **2019**, *31*, 89–94.

Manuscript received: November 8, 2023
Accepted manuscript online: December 29, 2023
Version of record online: ■■, ■■

© 2023 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

RESEARCH ARTICLE

The selective molecular recognition properties of two different calix[4]pyrrole receptors toward the *cis* isomer of acetylated lysine (*cis*-**Kac**) are investigated. The formation of a 1 : 1 complex is favoured for the aryl-extended receptor compared to the superaryl-extended featuring a deeper cavity for which the energetically not favourable desolvation of the aminoacidic group hampers the guest inclusion in the hydrophobic cavity.



Dr. M. Orlandini, Dr. A. Pedrini, D. Marchetti, Y. Li, Dr. G. Aragay, Prof. E. Dalcanale, Prof. P. Ballester, Prof. R. Pinalli**

1 – 8

Binding of Acetylated Lysine by Using a Water Soluble Aryl Extended Calix[4]pyrrole

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202303715 by Consorci De Serveis Universitaris De Catalunya, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License