Biophysical Chemistry 312 (2024) 107288

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: www.elsevier.com/locate/biophyschem

Isolated auto-citrullinated regions of PADI4 associate to the intact protein without altering their disordered conformation

José L. Neira [a,b,*], Bruno Rizzuti [b,c], Olga Abian [b,d,e,f], Adrian Velazquez-Campoy [b,d,e,f]

[a] IDIBE, Universidad Miguel Hernández, 03202 Elche, Alicante, Spain
[b] Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Universidad de Zaragoza, 50018 Zaragoza, Spain
[c] CNR-NANOTEC, SS Rende (CS), Department of Physics, University of Calabria, 87036 Rende, Italy
[d] Instituto de Investigación Sanitaria Aragón (IIS Aragón), Zaragoza, Spain
[e] Centro de Investigación Biomédica en Red en el Área Temática de Enfermedades Hepáticas y Digestivas (CIBERehd), 28029 Madrid, Spain
[f] Departamento de Bioquímica y Biología Molecular y Celular, Universidad de Zaragoza, 50009 Zaragoza, Spain

ARTICLE INFO

Keywords:
Citrullination
Disordered peptides
Fluorescence
Isothermal titration calorimetry
Kinetics
Molecular docking

ABSTRACT

PADI4 is one of the human isoforms of a group of enzymes intervening in the conversion of arginine to citrulline. It is involved in the development of several types of tumors, as well as other immunological illnesses, such as psoriasis, multiple sclerosis, or rheumatoid arthritis. PADI4 auto-citrullinates in several regions of its sequence, namely in correspondence of residues Arg205, Arg212, Arg218, and Arg383. We wanted to study whether the citrullinated moiety affects the conformation of nearby regions and its binding to intact PADI4. We designed two series of synthetic peptides comprising either the wild-type or the relative citrullinated versions of such regions – i.e., a first series of peptides comprising the first three arginines, and a second series comprising Arg383. We studied their conformational properties in isolation by using fluorescence, far-ultraviolet (UV) circular dichroism (CD), and 2D—^{1}H NMR. Furthermore, we characterized the binding of the wild-type and citrullinated peptides in the two series to the intact PADI4, by using isothermal titration calorimetry (ITC), fluorescence, and biolayer interferometry (BLI), as well as by molecular docking simulations. We observed that citrullination did not alter the local conformational propensities of the isolated peptides. Nevertheless, for all the peptides in the two series, citrullination slowed down the kinetic k_{off} rates of the binding reaction to PADI4, probably due to differences in electrostatic effects compared to the presence of arginine. The affinities of PADI4 for unmodified peptides were slightly larger than those of the corresponding citrullinated ones in the two series, but they were all within the same range, indicating that there were no relevant variations in the thermodynamics of binding due to sequence effects. These results highlight details of the self-citrullination of PADI4 and, more generally, of possible auto-catalytic mechanisms taking place in vivo for other citrullinating enzymes or, alternatively, in proteins undergoing citrullination passively.

1. Introduction

Peptidyl-arginine deiminases (PADI, EC 3.5.3.15) are enzymes catalyzing the post-translational modification (PTM) of polypeptide-bound arginine residues, known as citrullination, in the presence of Ca (II) ions. This PTM is an irreversible chemical reaction believed to significantly impact the molecular properties of a protein, as it causes loss of a positive charge and the possible onset of new protein-protein interactions [1]. Moreover, citrullination has important roles in human health and on several diseases, such as tumors, psoriasis, multiple sclerosis, and rheumatoid arthritis (RA) [2–6]. The importance of citrullination in proteins was unclear until the first decade of 2000s, when two important discoveries brought attention to this PTM [2–5]. First, histones can be citrullinated, thus raising the possibility that, similarly to many other PTMs (such as phosphorylation, acetylation, and methylation), this modification can be crucial in transcription or other

Abbreviations: BLI, biolayer interferometry; CD, circular dichroism; DIPSI, decoupling in the presence of scalar interactions; ITC, isothermal titration calorimetry; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; PADI, peptidyl-arginine deiminase; PTM, post-translational modification; TCEP, tris(2-carboxyethyl)phosphine; TOCSY, total correlation spectroscopy; UV, ultraviolet.

* Corresponding author at: IDIBE, Edificio Torregaitán, Universidad Miguel Hernández, Avda. del Ferrocarril s/n, 03202 Elche, Alicante, Spain.
E-mail address: jlneira@umh.es (J.L. Neira).

https://doi.org/10.1016/j.bpc.2024.107288
Received 17 January 2024; Received in revised form 19 June 2024; Accepted 26 June 2024
Available online 29 June 2024

nuclear events involving nucleosomes. And secondly, patients with RA have an excess of antibodies recognizing peptidyl-citrulline epitopes, and those are highly specific for both diagnosis and assessment of disease severity [7].

There are five human genes encoding PADI proteins: PADI1, PADI2, PADI3, PADI4 and PADI6 [8–13]. All of them have molecular masses ranging from 74 to 77 kDa for the monomer (i.e., 663–694 residues). Each *PADI* gene has a different pattern of expression depending on the cell type, tissue, cell differentiation stage, and physiological or pathological conditions, resulting in a specific localization of the protein expressed. For instance, PADI4 is usually located in cytoplasmic granules of inflammatory cells (eosinophils, neutrophils and macrophages), mammary gland cells, stem cells, and tumor cells, where it is highly expressed, either in the cytosol or in the nucleus. PADI4 is involved in gene transcription and immune system modulation, leading to cell inflammation and an immune response [14–18], and there is clear evidence that PADI4 intervenes in RA development [19]. In fact, the synovium fluid in RA patients contains numerous citrullinated proteins, such as vimentin, alpha-enolase, type II collagen, fibrinogens alpha and beta, and fibrin. PADI4 is also capable of auto-deiminating itself [20–22]. This self-modification of PADI4 could potentially regulate its own activity and conformational stability, as well as its interactions with other proteins, calcium-dependent function, and substrate specificity [23–25].

In this work, we selected two regions of PADI4 which appear to be citrullinated in RA [4,7,20], namely A^{204}RSEMDKVRVFQATRGK^{220} and G^{375}LKEFPIKRVMGPDFGYVTR^{394}, to study their affinity for intact PADI4 and conformational propensities, for both the isolated wild-type and citrullinated sequences. We are aware that there are many arginines in the PADI4 sequence, but we chose those peptide sequences because those residues were citrullinated in RA [4,7,20]. For the first region (ARSE peptides), we designed four peptides: the unmodified sequence, and three peptides with a citrulline in each position, having an Arg residue in the parent sequence; for the second region (GLE peptides), we designed two peptides: the wild-type, and that with a citrulline instead of Arg383. It could be thought that for the ARSE peptides the study of the double or triple citrullinated variants could be worth investigating, but our interest was the study of the effect of a single citrulline in the sequence, without any possible compensatory effects due to nearby citrullines; besides, we have observed in other short polypeptide chains that the introduction of two or more citrullines in the sequences yields to aggregation of the corresponding double or triple modified fragment (unpublished data). We used an array of *in vitro* biophysical and spectroscopic techniques, namely fluorescence, far-ultraviolet (UV) circular dichroism (CD), biolayer interferometry (BLI), nuclear magnetic resonance (NMR), and isothermal titration calorimetry (ITC), together with *in silico* docking calculations, to address the binding to PADI4 and the local conformational propensities of the arginine or citrulline residues in these isolated peptides. We observed that all the ARSE and GLE peptides were monomeric and disordered in aqueous solution, and that the citrullination only induced minor, local changes in the chemical shifts of the side-chains of the citrulline residues. Binding to PADI4 occurred with all of these peptides, as tested by fluorescence, BLI, ITC, and molecular docking. The affinity was more favorable (i.e., smaller K_D) for the unmodified peptides than for the citrullinated ones, as observed by fluorescence and ITC (in both ARSE and GLE series), although the affinities of all of them for PADI4 were within the same range. Furthermore, the kinetic dissociation rates of the citrullinated peptides in the binding reaction, when compared to the corresponding unmodified peptide, were slower in both ARSE and GLE series, and this effect seems to be more evident for the former peptides. The implications of our findings are that, for the first time: (i) the thermodynamics and kinetics of binding of citrullinated peptides to a PADI enzyme are measured; (ii) the conformational effects of a single citrullination on isolated polypeptide sequences are studied to see whether the position affects the observed conformations; and, (iii) there is an attempt to dissect out the functioning of PADI4 by understanding from a kinetic and thermodynamic point of view the binding of citrullinated peptides to the enzyme, and most importantly, to envisage how this information may be used for the design of peptide inhibitors.

2. Materials and methods

2.1. Materials

Imidazole, Trizma base, DNase, SIGMAFAST protease tablets, NaCl, Ni^{2+}-resin, 3-(trimethylsilyl) propionic acid-2,2,3,3-$^2H_4$-sodium salt (TSP), Amicon centrifugal devices with a molecular weight cut-off of 50 kDa and ultra-pure dioxane were from Sigma-Merck (Madrid, Spain). The β-mercaptoethanol was from BioRad (Madrid, Spain). Ampicillin and isopropyl-β-D-1-thiogalactopyranoside were obtained from Apollo Scientific (Stockport, UK). Triton X-100, TCEP and the SDS protein marker (PAGEmark Tricolor) were from VWR (Barcelona, Spain). The rest of the materials used were of analytical grade. Water was deionized and purified on a Millipore system.

2.2. Protein expression and purification

PADI4 is a homodimer, and it was purified in its dimeric form as previously described [6,26–28]. Protein concentration was determined by UV absorbance, employing an extinction coefficient at 280 nm estimated from the ten tryptophans and thirteen tyrosines *per* monomer (73,540 M^{-1} cm^{-1}) [29].

2.3. Peptide design

Peptides were synthesized and purified by GenScript (Amsterdam, Netherlands) with the corresponding N terminus acetylated and the C terminus amidated. The sequences were:

(a) ARSE peptide series: the wild-type peptide (P1): A^{204}RSEMDKVRVFQATRGK^{220}; the peptide with the first arginine modified to citrulline (indicated as an **X** in the following sequences) (P2): A^{204}**X**SEMDKVRVFQATRGK^{220}; the peptide with the second arginine modified to citrulline (P3): A^{204}RSEMDKV**X**VFQATRGK^{220}; and the last peptide in the series with the third arginine modified to citrulline (P4): A^{204}RSEMDKVRVFQAT**X**GK^{220}.We have not employed the double (three peptides) and triple citrulline modifications because of: (i) the potential aggregation propensity elicited by citrullination in other peptides (JLN, BR and AVC, unpublished results); and, (iii) the need to describe the influence of a single citrulline at a time in a short sequence, without interference of other nearby citrullines, to determine whether the position of the modified residue affects the observed conformations.

(b) GLE series: the wild-type peptide (P5) with the sequence G^{375}LKEFPIKRVMGPDFGYVTR^{394}; and the citrullinated peptide obtained by modifying the arginine in the central position (P6): G^{375}LKEFPIK**X**VMGPDFGYVTR^{394}.

Peptide concentration was determined by UV absorbance, employing an extinction coefficient at 280 nm estimated from the single phenylalanine (700 M^{-1} cm^{-1}) in the ARSE peptides, or the two phenylalanines and single tyrosine (2680 M^{-1} cm^{-1}) in the GLE peptides [29].

2.4. Fluorescence

2.4.1. Steady-state fluorescence

Fluorescence spectra were collected on a Cary Varian spectrofluorometer (Agilent, Santa Clara, CA, USA), interfaced with a Peltier unit. All experiments were carried out at 25 °C. Following the standard protocols used in our laboratories, the samples were prepared the day

before and left overnight at 5 °C; before experiments, samples were left for 1 h at 25 °C. A 1-cm-pathlength quartz cell (Hellma, Kruibeke, Belgium) was used. Concentration of PADI4 was 2.5 μM (in protomer units), and that of the corresponding peptide was 20 μM. Experiments were performed in 20 mM Tris buffer (pH 7.5), 5 mM TCEP, 150 mM NaCl and 5% glycerol. Polypeptides were excited at either 280 or 295 nm. The other experimental parameters for the acquisition of the fluorescence spectra have been described elsewhere [30]. Appropriate blank corrections were made in all spectra. Fluorescence experiments were repeated in triplicates with newly prepared samples. Variations of results among the experiments were lower than 10%.

2.4.2. Binding experiments with PADI4

For the titration of the ARSE or GLE peptides with PADI4, increasing amounts of the corresponding peptide, in the concentration range 0–30 μM, were added to a solution with a fixed concentration of PADI4 (2.0 μM in protomer units). Experiments were carried out in the same buffer used for the steady-state fluorescence experiments. The samples were excited at either 280 or 295 nm, and the rest of the experimental set-up was the same as described above. In all cases, the appropriate blank corrections were made by subtracting the signal obtained with the corresponding amounts of ARSE or GLE peptides, by using KaleidaGraph (Synergy Software, Reading, PA, USA). Spectra were corrected for inner-filter effects during fluorescence extinction [31]. The titrations were repeated three times for each peptide, using newly prepared samples. In the three cases, the variations in the results were lower than 10%.

The samples for the titration experiments were prepared the day before and left overnight at 5 °C; before the measurements, they were incubated for 1 h at 25 °C. The dissociation constant of the corresponding complex, K_D, was calculated by fitting the binding isotherm obtained from the variation of the fluorescence intensity as a function of peptide concentration to the general binding model, explicitly considering ligand depletion [32,33]:

$$F = F_0 + \frac{\Delta F_{max}}{2[PADI4]_T}\left([\text{peptide}]_T + [\text{PADI4}]_T + K_D\right) - \sqrt{\left(\left([\text{peptide}]_T + [\text{PADI4}]_T + K_D\right)^2 - 4[\text{peptide}]_T[\text{PADI4}]_T\right)}\,\text{F}$$
$$= \text{F}_0\frac{\Delta \text{F}_{\max}}{2[\text{PADI4}]_\text{T}}\left([\text{Impα3}-\text{species}]_\text{T} + [\text{PADI4}]_\text{T} + \text{K}_\text{d}\right) - \sqrt{\left(\left([\text{Impα3}-\text{species}]_\text{T} + [\text{PADI4}]_\text{T} + \text{K}_\text{d}\right)^2 - 4[\text{Impα3}-\text{species}]_\text{T}[\text{PADI4}]_\text{T}\right)}\text{F}$$
$$= \text{F}_0\frac{\Delta \text{F}_{\max}}{2[\text{PADI4}]_\text{T}}\left([\text{Impα3}-\text{species}]_\text{T} + [\text{PADI4}]_\text{T} + \text{K}_\text{d}\right) - \sqrt{\left(\left([\text{Impα3}-\text{species}]_\text{T} + [\text{PADI4}]_\text{T} + \text{K}_\text{d}\right)^2 - 4[\text{Impα3}-\text{species}]_\text{T}[\text{PADI4}]_\text{T}\right)} \quad (1)$$

where F is the measured fluorescence intensity at any particular concentration of each peptide, after subtraction of the spectrum containing the same concentration of the sole polypeptide; ΔF_{max} is the largest change in the fluorescence of PADI4 when all polypeptide molecules were forming the complex, compared to the fluorescence of each isolated chain; F_0 is the fluorescence intensity when no peptide was added; $[PADI4]_T$ is the constant, total concentration of PADI4 (2.0 μM in protomer units); and $[peptide]_T$ is that of each peptide from the ARSE or GLE series, which was varied during the titration. Fitting to the above equation was carried out by using KaleidaGraph.

2.5. Circular dichroism (CD)

Far-UV CD spectra were collected on a Jasco J810 spectropolarimeter (Jasco, Tokyo, Japan) with a thermostated cell holder and interfaced with a Peltier unit. The instrument was periodically calibrated with (+)-10-camphorsulfonic acid. A cell of pathlength 0.1 cm was used (Hellma, Kruibeke, Belgium). All spectra were corrected by subtracting the corresponding baseline. Concentration of each polypeptide (both PADI4 and the specific peptide added) and the buffers were the same used in the fluorescence experiments. Samples were prepared in the same way as in the fluorescence experiments.

Isothermal wavelength spectra of each isolated polypeptide and that of the complex were acquired as an average of 6 scans, at a scan speed of 50 nm/min, with a response time of 2 s and a band-width of 1 nm.

2.6. Nuclear Magnetic Resonance (NMR) spectroscopy

The NMR spectra of the isolated ARSE or GLE peptides were acquired at 10 °C on a Bruker Avance 500 MHz spectrometer (Bruker GmbH, Germany), equipped with a triple resonance probe and z-pulse field gradients. Spectra were processed with Bruker TopSpin 2.1 (Bruker GmbH, Karlsruhe, Germany). The temperature of the probe was calibrated with pure methanol [34]. All NMR experiments with GLE peptides were carried out in 100 mM sodium phosphate buffer (not corrected for isotope effects), at pH 7.0; those of ARSE peptides were carried out at pH 3.0 (glycine buffer), because the isolated peptides precipitated at pH 7.0 (both in phosphate and Tris buffer). Spectra were calibrated with TSP, by considering the pH-dependent changes of its chemical shifts [34].

2.6.1. 1D—^{1}H NMR spectra

A number of 64 scans were acquired with 16 K acquisition points for the homonuclear 1D—^{1}H NMR spectra of each isolated peptide at a concentration of 1.2 mM. Water signal was suppressed with the WATERGATE sequence [35]. The spectra were processed after zero-filling and apodization was carried out with an exponential window.

2.6.2. Translational diffusion NMR (DOSY)

The concentrations of ARSE or GLE peptides series in DOSY experiments was 100 μM, and 128 scans were acquired, in which the gradient strength was varied linearly. Measurements of the translational self-diffusion were performed with the pulsed-gradient spin-echo sequence in the presence of 100% D_2O. Details on the experimental conditions and the protocols of fitting of the resulting curves have been described elsewhere [36]. During the acquisition, the gradient strength was varied in sixteen linear steps between 2 and 95% of the total power of the gradient coil. Gradient strength was calibrated by using the value of the translational diffusion coefficient, D, for the residual proton water signal in a sample containing 100% D_2O, in a 5-mm tube [37]. In the experiments with the ARSE or GLE peptides, the gradient length was 2.5 ms; the time between the two pulse gradients in the pulse sequence was 250 ms; and the recovery delay between the bipolar gradients was 100 μs. The signals of the methyl groups between 0.8 and 1.0 ppm were used for peak integration, for all ARSE or GLE peptide series. A final concentration of 1% of dioxane, which was assumed to have a hydrodynamic radius R_h of 2.12 Å [37], was added to the solution to be used as a reference to measure the R_h value.

2.6.3. 2D—^{1}H NMR spectra

Two-dimensional NMR spectra of the ARSE and GLE peptides were acquired in each dimension in phase-sensitive mode by using the time-proportional phase incrementation technique [38] and a spectral

width of 5500 Hz. Peptide concentration was the same used in the 1D experiments and the pH values were those reported above. Standard TOCSY [39] (total correlation spectroscopy), and NOESY (nuclear Overhauser effect spectroscopy) [40], with mixing times of 80 and 250 ms, respectively, were performed by acquiring a data matrix size of 4096 × 256 points. The relaxation time in both experiments was 1 s. The DIPSI (decoupling in the presence of scalar interactions) spin-lock sequence [41] was used in the TOCSY experiments. A number of 88 scans were acquired per increment in the first dimension, and the residual water signal was removed by using the WATERGATE sequence [35]. NOESY spectra were collected with 128 scans per increment in the first dimension, using again the WATERGATE sequence [35]. Data were zero-filled, resolution-enhanced with a square sine-bell window function optimized for each spectrum, and baseline-corrected. The ^{1}H resonances were assigned by standard sequential assignment processes [34]. The chemical shift values of H_α protons in random-coil regions were obtained from comparison with the values of model polypeptides, corrected by neighboring residue effects [34,42,43].

2.7. Isothermal titration calorimetry (ITC)

Calorimetric titrations for testing the interaction of PADI4 with ARSE or GLE peptides were carried out in an Auto-iTC200 automated high-sensitivity calorimeter (MicroCal, Malvern-Panalytical, Madrid, Spain). Experiments were performed at 25 °C in 20 mM Tris buffer (pH 7.5), 5 mM TCEP, 150 mM NaCl and 5% glycerol at 25 °C. Peptide solution (50 μM) in the injection syringe was titrated into the PADI4 solution (5 μM) in the calorimetric cell. A series of 19 injections with 2 μL volume, 0.5 μL/s injection speed, and 150 s time spacing was programmed while maintaining a reference power of 10 μcal/s and a stirring speed of 750 rpm. The heat effect *per* injection was calculated by integration of the raw data of the thermal power after baseline correction, and the interaction isotherm curve (ligand-normalized heat effect *per* injection as a function of the molar ratio) was obtained by a non-linear least squares regression data analysis, applying a model that considers a single binding site to estimate the association constant, K_a, the interaction enthalpy, ΔH, and the stoichiometry of binding n (although, in practice, the last parameter usually reports the fraction of active protein in the sample within the calorimetric cell). The dissociation constant was calculated from the association constant: $K_D = 1/K_a$. The background injection heat (usually called "dilution heat", but reflecting any unspecific phenomenon such as titrant/solute dilution, buffer neutralization, temperature equilibration, or solution mechanical mixing) was accounted for by including an adjustable constant parameter in the fit. The data analysis was conducted in Origin 7.0 (OriginLab, Northampton, MA, USA) with user-defined fitting functions.

2.8. Biolayer interferometry (BLI)

2.8.1. Experimental design of BLI experiments

The kinetic association (k_{on}) and dissociation (k_{off}) rate constants of the binding of the ARSE or GLE peptides to PADI4 were determined by using a BLItz system (ForteBio Pall, Barcelona, Spain) [44]. The buffer used in the experiments was that recommended by the manufacturer. As PADI4 had a His-tag, it was immobilized on His-tag biosensors (Forte Bio) at the concentration of 0.45 μM. The peptide concentration was varied in the range from 1 to 12 μM during the association step. The general schemes of the protein-peptide association/dissociation reactions in the BLItz system were: 30 s of acquisition of the initial baseline with the 10 × kinetics buffer; 120 s of loading PADI4 into the biosensor; 30 s of baseline acquisition with the 10 × kinetics buffer; 120 s of association of the peptide to the biosensor (which had been previously loaded with PADI4); and 120 s of dissociation of the peptide from the biosensor.

2.8.2. Fitting of the sensorgrams

Fittings of the sensorgrams were carried out by using KaleidaGraph. The interferometry response during the association step (measured in response units, RU), $R(t)$, and the binding rate, $dR(t)/dt$, were used to evaluate the kinetics of the formation of the PADI4/peptide complex, according to:

$$\frac{dR(t)}{dt} = k_{on}[\text{peptide}](R_{max} - R(t)) - k_{off}R(t) \quad (2)$$

where R_{max} is proportional to the total concentration of biosensor-bound PADI4; and [peptide] represents the corresponding concentration of a peptide from the ARSE or GLE peptide series.

In Eq. (2), $R(t)$ is given by:

$$R(t) = R_{eq} - R_{eq}e^{(-k_{obs}\ (t-t_0))} \quad (3)$$

where R_{eq} is the steady-state (or equilibrium) response obtained at infinite time when $dR(t)/dt = 0$, and $t_0 = 180$ s is the time at which the association step between biosensor-immobilized PADI4 and the peptide in the solution started. Since we observed a slope at the largest observation times, we fitted the experimentally obtained $R(t)$ under any condition as:

$$R(t) = R_{eq} - R_{eq}e^{(-k_{obs}\ (t-t_0))} - R_{eq}'(t - t_0) \quad (4)$$

The k_{obs} was used for the pseudo-first order plots, where its value is given by:

$$k_{obs} = k_{on}\ [\text{peptide}] + k_{off} \quad (5)$$

The dissociation process was always fitted to a single exponential, with $R(t)$ given by:

$$R(t) = R_1\ e^{(-k_{off}(t-t_0))} \quad (6)$$

where $t_0 = 300$ s is the time at which the dissociation of peptide from the biosensor-bound PADI4 started in our experimental set-up, and R_1 is the response level when dissociation starts.

2.9. Molecular modeling and simulation

Simulations were performed to study the binding to PADI4 of our peptides by using the docking engine AutoDock Vina, version 1.1.2 [45]. The peptides of the ARSE and GLE series encompassed 17 and 20 residues each, respectively, with cappings consisting of N-terminal acetylation and C-terminal amidation. The structure of PADI4 was built in dimeric form as previously described [6,26,28], and included two monomers of 663 amino acid residues each. The absence of any clear indication on the possible anchoring location of the peptides suggested the use of blind docking simulations on the whole quaternary structure of PADI4. Furthermore, due to the completely disordered structure of the peptides earlier evidenced in our experimental findings (NMR and far-UV CD), no rotation of their dihedral angles could be constrained, resulting in a large number of degrees of freedom (75 and 79 for the ARSE and GLE peptide series, respectively).

To reduce the computational complexity of the simulation to a treatable level, we adopted a protocol we had previously employed for intrinsically disordered protein regions [46,47], including the two nuclear localization sequences of PADI4 [48]. This protocol consists in considering a sequence of 6-residue fragments, each shifted by two residues with respect to the previous one, and covering the whole sequence of each peptide. To avoid introducing fictitious charges at the fragments termini, N-terminal acetylation and C-terminal *N*-methyl capping were added. The use of 6-residue peptide fragments reduced the number of degrees of freedom of each sequence portion to a maximum of 28 rotatable bonds, which is within the limit (≤ 32 torsions) considered reliable for using the AutoDock Vina algorithm [45]. Furthermore, the structure of PADI4 was considered as rigid, and both the protein and

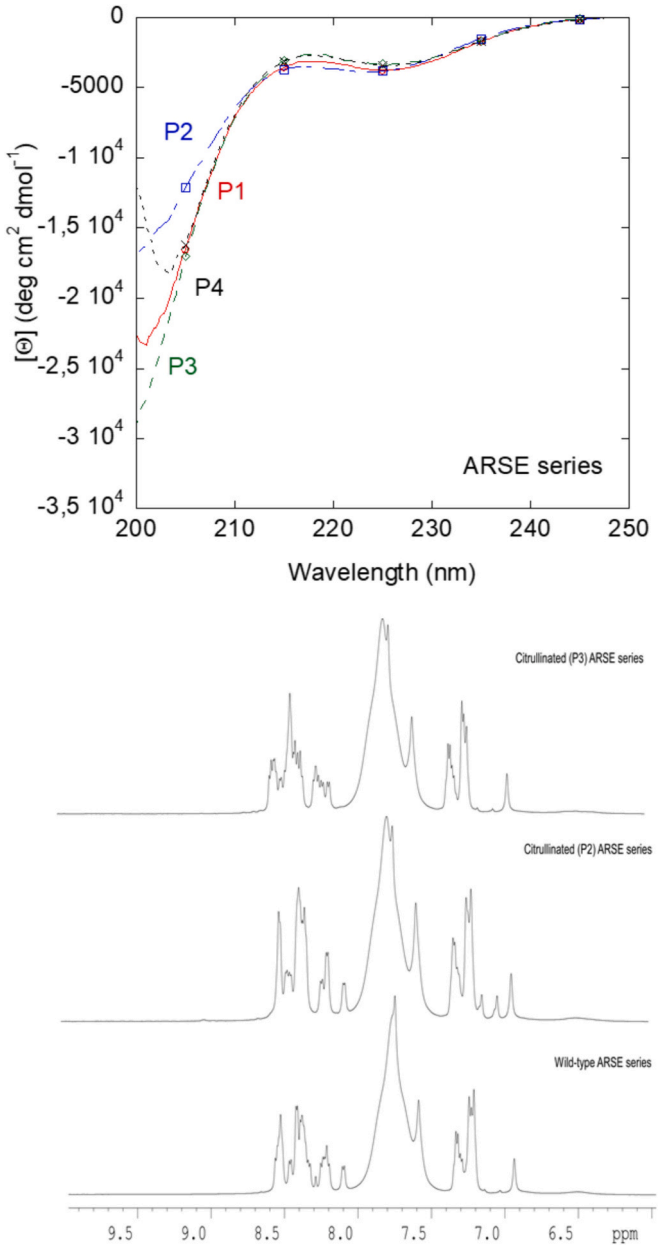

Fig. 1. Conformational features of the isolated ARSE peptides: Far-UV CD (top panel) and 1D—^{1}H NMR (amide region) spectra of selected ARSE peptides (bottom panel). P1 represents the wild-type peptide of the ARSE series. The sequences of the peptides (where X indicates a citrulline in the following sequences) are: P1: A^{204}RSEMDKVRVFQATRGK^{220}; P2: A^{204}XSEMDKVRVFQATRGK^{220}; P3: A^{204}RSEMDKVXVFQATRGK^{220}; and P4: A^{204}RSEMDKVRVFQATXGK^{220}.

each peptide fragment were described in simulation with a united-atom approximation, by modeling apolar hydrogens as implicit.

3. Results

3.1. Conformational features and PADI4-binding of the ARSE peptide series

3.1.1. Conformational features of isolated ARSE peptides

As the ARSE peptide series do not have any fluorescent residue, this technique cannot be used to gain structural information on them. We started to investigate their conformational properties by acquiring the far-UV CD spectra of the isolated peptides (Fig. 1, top). All the ARSE peptides showed an intense minimum at ~200 nm, indicating that they possessed mostly a random-coil conformation. All spectra had a small shoulder at 222 nm, which could be attributed to the presence of the sole aromatic residue in the sequence, Phe214 [49–51]. The disordered character of the ARSE series was further confirmed by their 1D—^{1}H NMR spectrum (Fig. 1 bottom), which showed a clustering of the signals of all the amide protons between 8.0 and 8.6 ppm, whereas those of the methyl protons were observed between 0.8 and 1.0 ppm. All these values for those kinds of protons are typical of disordered chains [34].

The isolated peptides were monomeric, as concluded from the *D* values measured by the DOSY experiment, ranging from $(1.120 \pm 0.006) \times 10^{-6}$ to $(1.08 \pm 0.04) \times 10^{-6}$ cm^2 s^{-1}. The estimated R_h values (obtained from the comparison with the *D* of dioxane) were around 11.7 ± 0.8 Å in all cases. This value of R_h was similar to that one obtained theoretically [52] for a random-coil polypeptide with the same molecular weight (2020.3 Da), i.e., 12 ± 2 Å.

To further confirm the disordered nature of the peptides of the ARSE series, we also carried out homonuclear 2D—^{1}H NMR experiments (Tables ST3-ST6). All the peptides were found to be mainly disordered in solution, as suggested by two lines of evidence, further confirming the results from far-UV CD (Fig. 1, top) and 1D—^{1}H NMR spectra (Fig. 1, bottom). First, the sequence-corrected conformational shifts (Δδ) of H_α protons [34,42,43] were within the commonly accepted range for random-coil peptides (Δδ ≤ 0.1 ppm). And second, no long- or medium-range NOEs were detected, but only sequential ones (i.e., αN (*i*, *i* + 1) and βN (*i*, *i* + 1)) (Fig. 2, top).

As a complementary finding from this work, we also identified the values of the chemical shifts of protons of the side-chain of citrulline (provided in Tables ST1-ST6). We note that, as far as we know, this is the first example of chemical shifts of protons for the side-chain of citrulline reported in the literature. The changes in chemical shifts in all peptides were all located around the side-chain corresponding to arginines in the wild-type sequence, whereas the rest of the amino acids were not affected by the modification.

Taken together, all the experimental techniques concurred to indicate that the isolated ARSE peptides were monomeric (i.e., no significant self-association was observed) and disordered in aqueous solution.

3.1.2. Binding of ARSE peptides to PADI4

We measured the affinity of ARSE peptides for PADI4 *in vitro*, by using three techniques. We carried out fluorescence emission (noting that the only fluorescent species in the complex was PADI4), BLI, and ITC experiments to quantitatively measure the thermodynamic binding parameters.

The fluorescence results provided a K_D value for any of the citrullinated peptides of 3 ± 1 μM (Table 1, Fig. 3 A) but, interestingly, the wild-type peptide did not yield any binding curve. This observation suggests either that peptide binding did not happen within that concentration range or, alternatively, that the changes induced on PADI4 tryptophans' fluorescence due to variations in their microenvironments were not indicative of peptide binding.

We also used ITC to assess whether the binding took place and, if so, to determine the thermodynamic binding parameters. An exothermic interaction was observed (Fig. 3 B), with a favorable enthalpic contribution, accompanied by an unfavorable or marginally favorable entropic contribution to the Gibbs energy of binding (Table 1). In all cases, the dissociation constant, K_D, was in the low micromolar range, exhibiting the highest affinity for the unmodified peptide (K_D = 4 μM for the wild-type sequence, compared to 7–10 μM for the citrullinated ones). The fact that we could measure the binding by ITC suggests that fluorescence was spectroscopically silent to binding for the wild-type peptide.

The results from BLI (Table 2), indicated that the dissociation constants, obtained from the ratio of k_{off} and k_{on}, were similar to those obtained by fluorescence (for the citrullinated peptides), thus suggesting

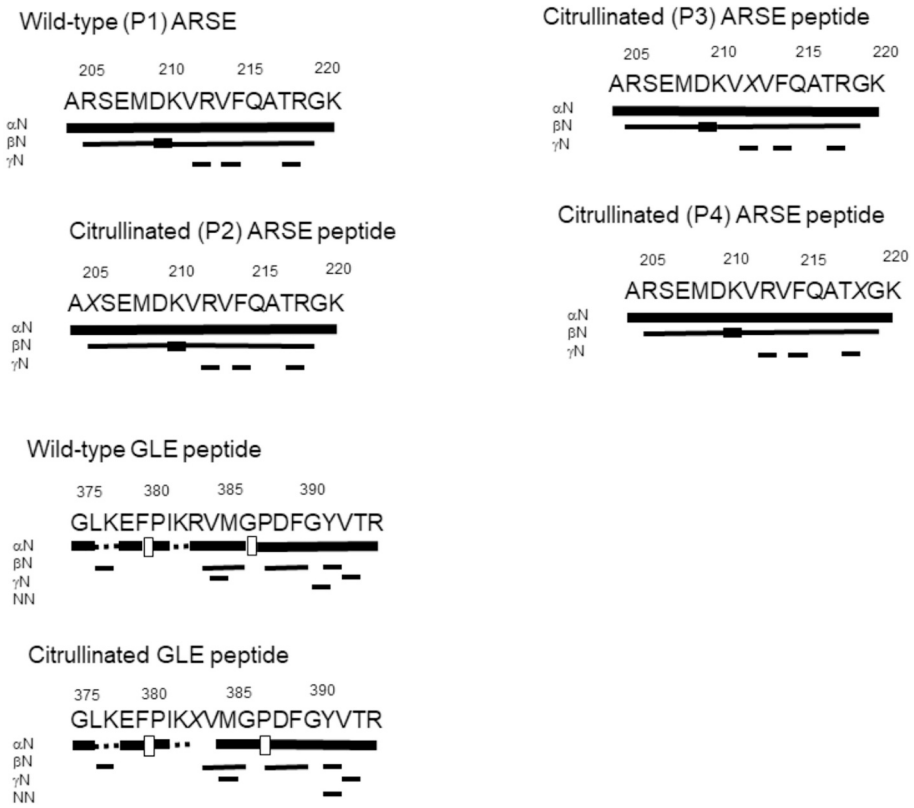

Fig. 2. NOE diagrams of the ARSE and GLE peptide series in aqueous solution: In the top panel, the NOE diagrams of the ARSE series are indicated; in the bottom panel, the NOE diagrams of GLE peptide series are shown. NOEs are classified as strong, medium or weak, and represented by the height of the bar underneath the sequence; signal intensity was judged from the NOESY experiments. The dotted lines indicate NOE contacts that could not be unambiguously assigned. A white square indicates an αδ(i,i + 1) NOEs involving a proline residue. The modified arginine appears as "*X*" in each diagram.

Table 1
Affinity (equilibrium) measurements of ARSE and GLE peptide series as determined by ITC and fluorescence (at 25 °C).

	ITC[a]		Fluorescence[b]
Peptide	K_D (μM)	ΔH (kcal/mol)	
P1 (wild-type)	4.0 ± 0.6	−17.1 ± 0.8	Not determined
P2 (citrullinated)	6.6 ± 0.9	−11.8 ± 0.5	2.8 ± 0.4
P3 (citrullinated)	7.3 ± 0.9	−5.4 ± 0.4	3 ± 1
P4 (citrullinated)	10 ± 2	−15.5 ± 0.6	3 ± 1
GLE (wild-type)	1.4 ± 0.3	−6.8 ± 0.5	Not determined
GLE (citrullinated)	6.8 ± 0.8	−12.0 ± 0.6	13 ± 10

[a] Errors are from fitting to the binding isotherm.
[b] Errors are from fitting to Eq. (1). For both wild-type peptides no titration was observed by fluorescence.

that the two-state kinetic model assumed for the binding reaction was correct. However, both kinetic rate constants were larger for the unmodified peptide than for the citrullinated ones, suggesting that citrullination slowed down the formation and dissociation of the peptide-PADI4 complex. Among the three citrullinated peptides, both kinetic rates (k_{on} and k_{off}) were within the same range. The fact that we could measure the binding by BLI further supports that fluorescence was spectroscopically silent to binding for the wild-type peptide.

3.2. Conformational features and PADI4-binding of the GLE peptide series

3.2.1. Conformational features of isolated GLE peptides

Because the GLE peptides have a tyrosine residue, it was possible to acquire the fluorescence spectrum of each isolated peptide (Fig. S1 B). The spectra of both peptides had a maximum around 308 nm, and therefore the presence of a citrullinated residue did not alter the fluorescence of the single tyrosine. The far-UV CD spectra of both peptides of the GLE series were similar to those of the ARSE series, with global minima around 200 nm (Fig. S1 A), indicating that also these peptides were disordered in solution. The disordered character of both GLE peptides was further confirmed by the 1D—^{1}H NMR spectra (Fig. S1 C), which showed a clustering of the signals of all the amide protons between 8.0 and 8.6 ppm, whereas the methyl protons were observed between 0.8 and 1.0 ppm. All these values are typical of disordered polypeptide chains [34].

Both isolated peptides were monomeric, as concluded from the value of D measured by the DOSY experiments, and the R_h value (obtained from the comparison with the translational diffusion coefficient of dioxane) were estimated. For the wild-type peptide, the D was (1.09 ± 0.01) × 10^{-6} cm^2 s^{-1} and the R_h was 12.25 Å; for the citrullinated peptide, the D was (1.101 ± 0.008) × 10^{-6} cm^2 s^{-1} and the R_h was 12.10 Å. The value of R_h was similar to that obtained theoretically for a random-coil polypeptide [52] with the same molecular weight (2310.74 Da): 13 ± 2 Å.

To further confirm the disordered nature of the GLE peptides, we also carried out homonuclear 2D—^{1}H NMR experiments (Tables ST1-ST2). All the peptides were mainly disordered in solution, as suggested by two lines of evidence, further confirming the results from far-UV CD (Fig. S1 A) and 1D—^{1}H NMR spectra (Fig. S1 C). First, the sequence-corrected conformational shifts (Δδ) of $H_α$ protons [34,42,43] for the assigned residues were within the commonly accepted range for random-coil peptides (Δδ ≤ 0.1 ppm). And second, no long- or medium-range NOEs were detected, but only sequential ones (Fig. 2, bottom), as it had already been observed for the ARSE peptides (Fig. 2, top).

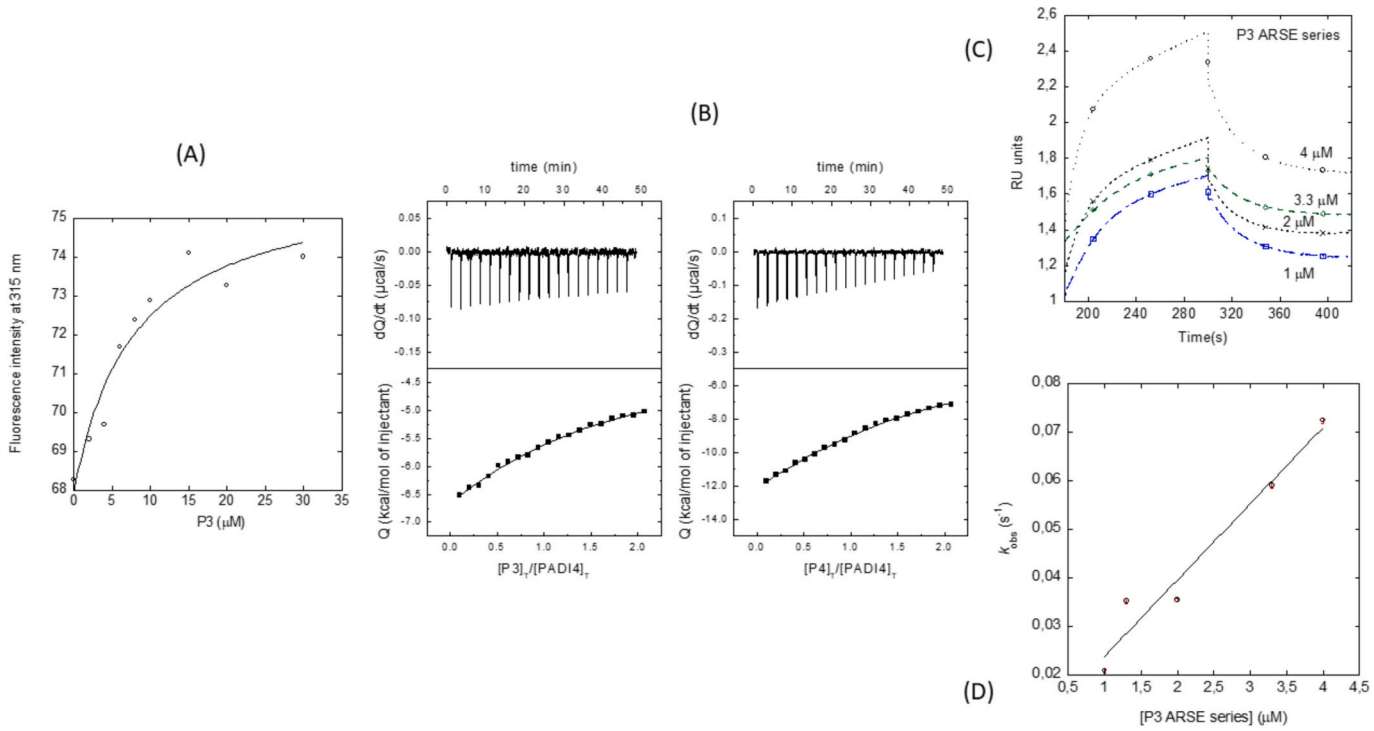

Fig. 3. Binding of ARSE peptides to PADI4 as monitored by biophysical probes: (A) Titration curve monitoring the changes in the fluorescence at 315 nm when peptide P3 was added to PADI4. The fluorescence intensity is the relative signal after removal of the corresponding blank. The line throughout the data is the fitting to Eq. (1). (B) Calorimetric titrations for PADI4 binding to selected peptides. Upper panels show the thermograms (thermal power as a function of time), and lower panels show the binding isotherms (ligand-normalized heat effects *per* injection, as a function of the molar ratio in the calorimetric cell). Continuous lines correspond to the fitting curves according to a single-ligand binding site interaction model. (C) Sensorgrams of peptide P3 at different concentration, in the assays with PADI4 immobilized on the sensor. (D) Pseudo-first order plot of the binding of peptide P3 to PADI4 (Eq. 5). The error bars are fitting errors from fitting sensorgrams to Eq. (4). Experiments were carried out at 25 °C.

Table 2
Kinetic parameters of the binding reaction of the ARSE and GLE peptide series with intact PADI4 (at 25 °C).

Peptide series	k_{on} (μM^{-1} s^{-1})[a]	k_{off} (s^{-1})[a]	K_D (μM) (= k_{off} / k_{on})[b]
P1 (wild-type)	0.80 ± 0.08	0.15 ± 0.01	2 ± 0.2
P2 (citrullinated)	0.0025 ± 0.0006	0.019 ± 0.007	7 ± 3
P3 (citrullinated)	0.0050 ± 0.0006	0.008 ± 0.004	1.6 ± 0.8
P4 (citrullinated)	0.008 ± 0.002	0.016 ± 0.008	2 ± 1
GLE (wild-type)	0.024 ± 0.002	0.055 ± 0.002	2.3 ± 0.2
GLE (citrullinated)	0.017 ± 0.004	0.04 ± 0.02	2 ± 1

[a] Errors are from fitting to Eq. (5).
[b] Errors are calculated as propagation errors.

Taken together, all the experimental techniques concurred to indicate that the isolated GLE peptides were monomeric (i.e., with no significant self-association) and disordered in aqueous solution.

3.2.2. *Binding of GLE peptides to PADI4*

We measured *in vitro* the affinity of GLE peptides for PADI4, by following a three-part experimental approach. We carried out fluorescence, BLI, and ITC experiments to quantitatively measure the thermodynamic parameters of such binding.

As it happens with the ARSE series, the wild-type peptide did not yield any binding curve in the fluorescence experiments, suggesting that either binding did not happen within that concentration range or, alternatively, the binding was spectroscopically silent to fluorescence (Fig. 4 A). Furthermore, the citrullinated peptide yielded an affinity of 13 ± 10 μM, which was similar, within the experimental uncertainty, to those determined for the ARSE peptides. We also used ITC to calculate the thermodynamic binding parameters (Fig. 4 B). The results indicated a similar thermodynamic profile as that of ARSE peptides: the interaction was again exothermic (Table 1), i.e., with a favorable enthalpic contribution, accompanied by an unfavorable (or, at best, marginally favorable) entropic contribution to the Gibbs energy of binding. Similarly, the dissociation constant K_D was in the low micromolar range, with the highest affinity found for the unmodified peptide (K_D = 1.4 μM, compared to 6.8 μM for the citrullinated peptide).

The results from BLI measurements (Table 2, Fig. 4 C) indicated that the affinity constant obtained from the ratio of the k_{off} and k_{on} was similar, within the error, to that obtained by fluorescence for the citrullinated peptide, thus suggesting that the kinetic two-state approach is also correct for this peptide series. As it had been observed for the ARSE peptides, both kinetic rate constants were slightly faster for the unmodified peptide compared to those of the citrullinated peptide, although the effect was not as large as that observed for the ARSE peptides. Thus, citrullination also slowed down the formation and dissociation of the corresponding peptide-PADI4 complex. Comparison of the kinetic rates among the two series of peptides indicated that the values were different in the two polypeptides.

3.3. *Docking of peptide fragments to PADI4*

3.3.1. *Binding affinity towards PADI4*

Simulation techniques are often used to investigate the binding of peptides to a target protein. Unfortunately, computational methods are more reliable when the binding complex is structurally well defined, and become progressively more complex when the molecular system requires an extensive conformational sampling. Furthermore, whereas kinetics effects can be captured for the binding of small compounds, they are difficult to obtain for ligands with a large number of degrees of freedom.

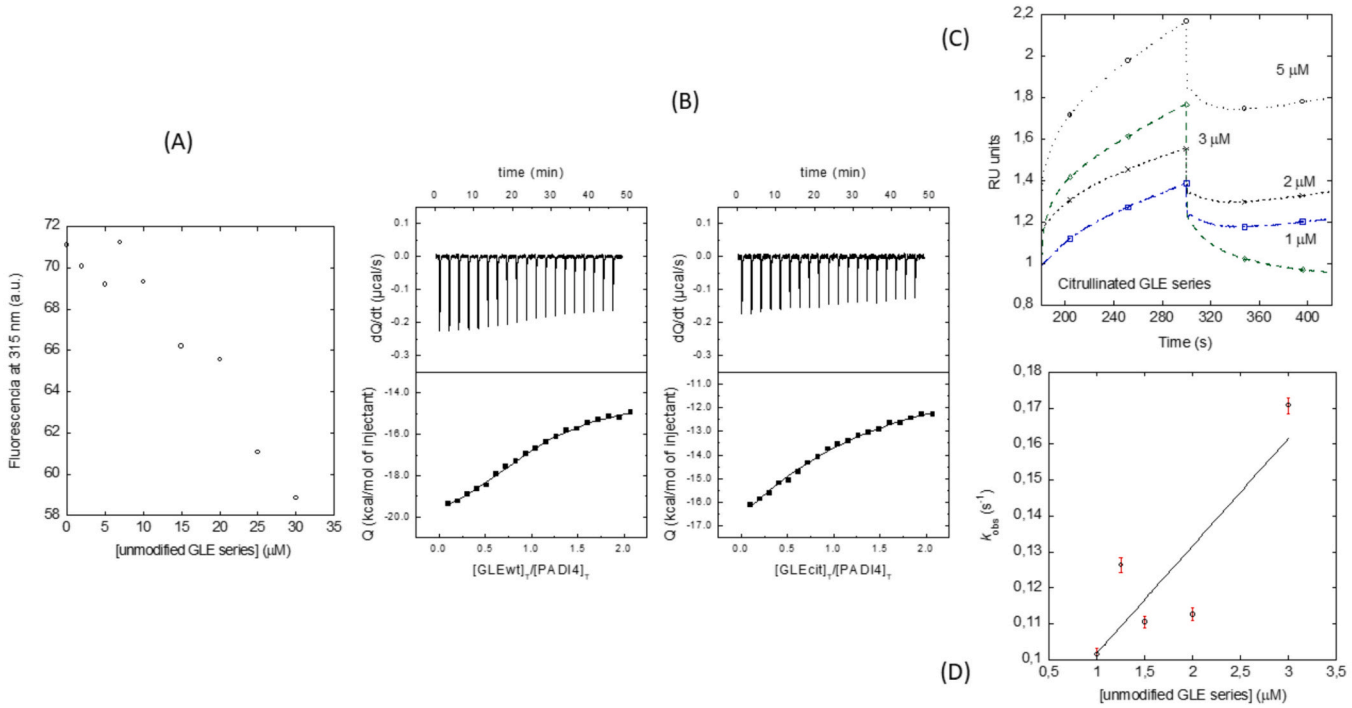

Fig. 4. Binding of GLE peptide series to PADI4 as monitored by biophysical probes: (A) Titration curve monitoring the changes in the fluorescence at 315 nm when the wild-type GLE peptide was added to PADI4. The fluorescence intensity on the y-axis is the relative signal after removal of the corresponding blank. (B) Calorimetric titrations for PADI4 binding to both peptides of the GLE series. Upper panels show the thermograms (thermal power as a function of time), and lower panels show the binding isotherms (ligand-normalized heat effects per injection, as a function of the molar ratio in the calorimetric cell). Continuous lines correspond to the fitting curves according to a single-ligand binding site interaction model (GLEwt indicates the unmodified peptide and GLEcit the citrullinated peptide). (C) Sensorgrams at different concentrations of the citrullinated GLE peptide, in the assays with PADI4-immobilized on the sensor. (D) Pseudo-first order plot of the binding of the wild-type GLE peptide (Eq. 5). The error bars are fitting errors from fitting sensorgrams to Eq. (4). Experiments were carried out at 25 °C.

In our case, the relatively large size of the peptides considered (17 and 20 amino acid residues for the ARSE and GLE peptide series, respectively), together with their disordered conformations detected in our *in vitro* experiments (NMR and CD), precluded the use of computationally-demanding simulation techniques to investigate the binding to PADI4. For these reasons, we used molecular docking, which is one of the simplest modeling techniques to assess the binding of a guest ligand to a host macromolecule. Furthermore, to overcome the relatively large number of rotatable dihedral angles in our peptides (75 and 79 degrees of freedom for each ARSE and GLE peptide, respectively), we followed a protocol we had already largely employed to study the binding of disordered peptide sequences to a target protein [47,53]. This methodology has also previously applied by using as docking guest some sequence fragments of PADI4 protein [48], or as docking host its full-length homodimer [26]. Accordingly, the association of 6-residue fragments belonging to each of the wild-type or citrullinated peptides to PADI4 was simulated, up to encompass the whole sequence for each peptide.

Fig. 5 reports the binding energy calculated for the ARSE and GLE 6-residue fragments, as a function of their residue number. Although not directly measuring the binding energy of the full-length peptides, these values provide information on the affinity in the anchoring of each of their sequence segment towards PADI4. The results showed a binding affinity in between about −7 and − 8 kcal/mol for each peptide fragment, which is sufficient to drive the binding. Furthermore, only minor variations were observed among the wild-type and citrullinated species in both series of peptides. As an example, in the case of the ARSE peptides (Fig. 5, top), the largest difference compared to the unmodified sequence was observed for the peptide citrullinated at position 212 in the sequence; however, the binding score was only 0.4 kcal/mol more favorable with respect to the wild-type species. This value is comparable to the expected uncertainty (0.2–0.3 kcal/mol) in the scoring procedure due to the stochastic search of AutoDock Vina [45], and below the energy due to thermal fluctuations ($R{\cdot}T_R \approx 0.6$ kcal/mol, where R is the gas constant and T_R is the room temperature). These observations suggest that differences in the binding due to citrullination are negligible from a thermodynamic point of view, in agreement with our fluorescence, BLI, and ITC experiments (Tables 1 and 2), all indicating that the dissociation constant of the wild-type and citrullinated species are similar, and within the low micromolar range in all cases.

3.3.2. *Anchoring location on the PADI4 surface*

The binding locations of all the peptide fragments (either unmodified or citrullinated) were also mapped on the surface of PADI4, as shown in Fig. 6. The most favorable docking poses were observed to cluster in four locations, which corresponded to two couples of symmetric hot-spot patches due to the head-to-tail quaternary structure of the PADI4 homodimer. Some of the fragments were bound to the immunoglobulin-like subdomain closer to the central region of PADI4, whereas others were bound to the distal immunoglobulin-like subdomain. This observation suggests that all our full-length ARSE and GLE peptides tend to associate in proximity of the interface of the two homodimers of PADI4, although possibly in a dynamic fashion that cannot be captured by docking experiments.

Docking poses of fragments in more solvent-exposed PADI4 regions, including the catalytic site, were occasionally observed with a less favorable binding affinity for both unmodified and citrullinated peptides, and could correspond to non-specific or metastable contacts of the peptides with the external protein surface, In particular, binding modes of arginine-containing fragments in the active site may favor the transient catalytic process of citrullination, although the concomitant presence of Ca(II) ions would also be necessary for this reaction to take place.

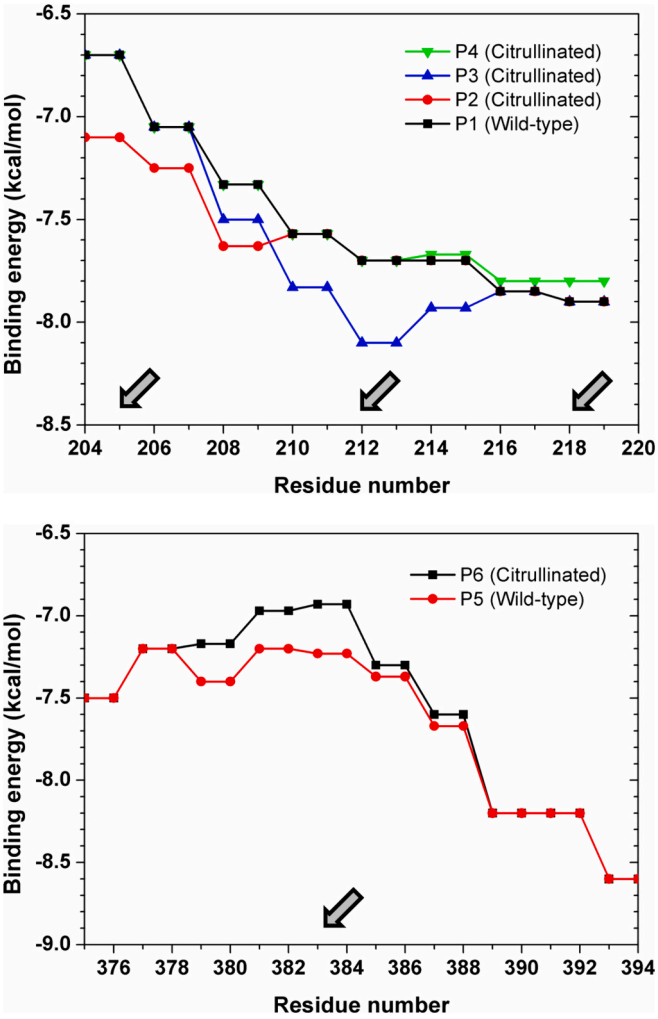

Fig. 5. Binding energy of the peptides to PADI4 as a function of their residue number: Affinities were calculated as a function of the sequence position, by averaging the docking scores obtained for all the fragments containing each residue in a given position. (Top) Peptides of the ARSE series; (bottom) peptides of the GLE series. Positions in the sequence that correspond to either arginine or citrulline are indicated with an arrow.

Remarkably, in our blind docking procedure (i.e., without the use of constraints guiding the binding to a selected target region), some docking poses of peptides fragments with a favorable binding affinity were observed to bind in the active site of PADI4 with a catalytic-competent geometry (see an example in Fig. S2), with the guanidine moiety of the citrullinatable arginine side-chain in direct contact (distance <5 Å) with the reactive sulfur atom of the catalytic residue Cys645.

It is interesting to note that docking poses in the active site occurred more frequently for unmodified rather than citrullinated sequences, due to a more favorable binding affinity of the former type of fragments (0.5 kcal/mol, on average) compared to the latter. This finding is not entirely surprising, because the existence of anchoring modes in the active site of PADI4 for a short sequence containing citrullination-prone arginines is a necessary prerequisite for these residues to be actually citrullinatable. Therefore, although unmodified peptides do not appear to target preferentially the active site, our docking simulations confirm that this binding location is only marginally less favorable (by about 0.5 kcal/mol) than the PADI4 homodimer interface, and cannot be entirely excluded as an anchoring hot-spot.

3.3.3. *Significance in the context of PADI4 auto-citrullination*

The computational techniques have a limited capacity to describe 'fuzzy' systems such as the ARSE and GLE peptides, which are disordered even in complex with PADI4. Nevertheless, in spite of the simplification consisting in the use of peptide fragments in our docking simulations, and although the results are not sufficient to model the structure of any full-length peptide bound to PADI4, two key conclusions can be summarized. First, the affinity of any fragment is sufficiently favorable to drive the binding (Fig. 5), confirming that enthalpic contributions are a major determinant of the peptide association to PADI4 (Table 1). Second, because any fragment (either unmodified or citrullinated) binds to the crevice of the PADI4 homodimer (Fig. 6), it is reasonable to assume that they all target the same crevice as the final binding location. Other regions showed a marginally lower binding affinity, and they might be possible alternative binding locations of the full-length peptides, for instance, as metastable contacts under the experimental conditions we have explored (or perhaps, as main binding sites under different conditions, such as different pH or ionic strength).

Interestingly, just as some conclusions about the peptide fragments could be generalized to the full-length peptides, it may also be possible to infer some further deductions about the potential significance of our findings when the peptides are part of the structure of the intact protein and, therefore, in the context of the possibility of auto-citrullination of PADI4. As shown in Fig. 7, although the sequence portions corresponding to our synthetic peptides are largely unstructured in the crystallographic structure of the intact PADI4 heterodimer, both of them also include a β-strand portion. This indicates that the conformational properties of these regions are not the same when they are within the intact PADI4 and in isolation. In addition, the binding properties of these regions are also not necessarily the same for intact PADI4 compared to the ARSE and GLE peptides. In particular, these regions cannot bind to the interface of another PADI4 dimer when they belong to an intact PADI4 dimer; in fact, this is sterically impossible, as judged from a visual inspection of the PADI4 structure (Fig. 7). More importantly, it can also be indirectly concluded from the fact that, if it were possible, this would lead to a self-association of multiple PADI4 dimers, which is not observed experimentally. In contrast, the catalytic site of PADI4 is more exposed to the solvent, and there is no steric hindrance to access it even when the two peptide sequences are part of another homodimer of PADI4. Thus, the auto-citrullination of PADI4 cannot be ruled out, and indeed it can be suggested as a possibility due to the fact that the unmodified fragments target the catalytic sites (Fig. S2) with a slightly less favorable binding affinity compared to the homodimer crevice.

It is worth noting that the auto-citrullination of PADI4 does not depend either on the necessity that the final fate of the two peptide sequences is to bind to the homodimer crevice after citrullination. This happens for the isolated peptides, but it should not occur for intact PADI4, for the same reasons stated above: (1) it is sterically impossible; and (2) it would imply a polymerization (self-association) reaction of citrullinated PADI4.

4. Discussion

Knowledge of the conformation of peptidyl substrates of PADI proteins is important for trying to find more efficient peptide inhibitors, or even more sensitive clinical assays and pharmacological treatments. For instance, the use of cyclic citrullinated peptides in an enzyme-linked immunosorbent assay (ELISA) of auto-antibodies against citrullinated proteins from RA patients results in a more sensitive assay than that performed by using linear citrullinated peptides [54]. On the other hand, a deeper knowledge of the structure of peptidyl substrates is equally useful, as conformational switches may take place as part of basic functional molecular mechanisms. For instance, the disordered N-terminal fragment of histones, another substrate of PADI4, acquires a β-turn-like conformation when bound to PADI4 [55]. In either case, our peptides represent an example of a model system to investigate the

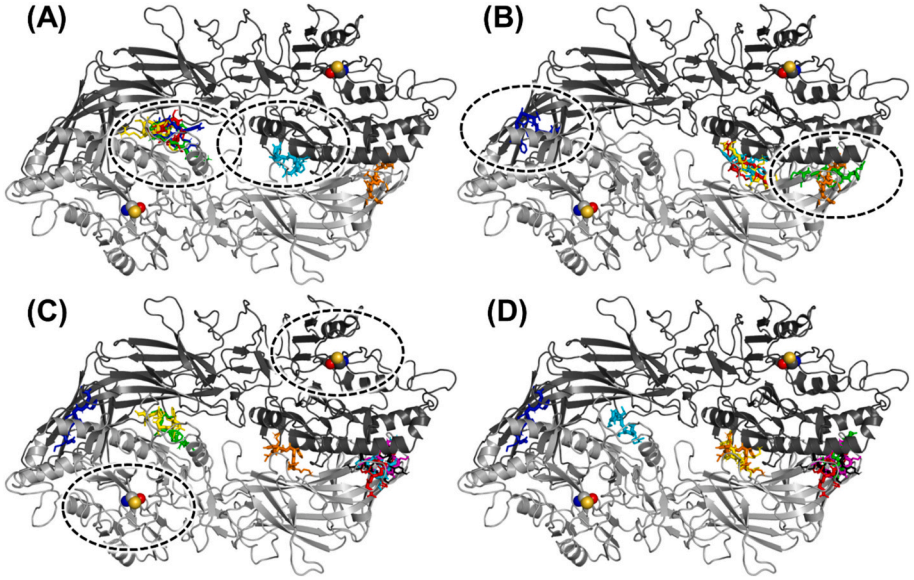

Fig. 6. Preferred binding location to PADI4 of the peptide fragments, as predicted in docking simulations: Protein monomers are in slightly different colors (dark and light gray), and the catalytic residue Cys645 is explicit shown (van der Waals representation). The 6-residue fragments are shifted by 2 residues from each other, up to encompass either peptide series, and they follow a rainbow colour scheme from the N to the C terminus (red → orange → yellow → green → cyan → blue → magenta → black; the last two colors are necessary only for the GLE peptides). Fragments are from (A) wild-type (P1) ARSE peptide; (B) citrullinated (P2-P4) ARSE peptides; (C) wild-type (P5) GLE peptide; (D) citrullinated (P6) GLE peptide. Note that some fragments are common to both unmodified and citrullinated peptides. The symmetric couple of hot-spot patches of PADI4 are indicated for the central immunoglobulin-like subdomain, the distal one, and the catalytic site (shown as circles in A, B, and C, respectively). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

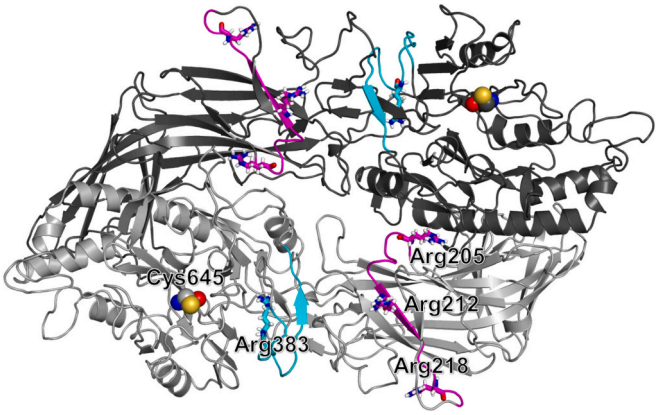

Fig. 7. Location of the two peptide sequences in the structure of the intact PADI4 heterodimer: The peptide P1 with sequence A^{204}RSEMDKVRVF-QATRGK^{220} (magenta) and P5 with sequence G^{375}LKEFPIKRVMGPDF-GYVTR^{394} (cyan) are shown in each monomer of PADI4, with their arginine residues labeled. The catalytic residues Cys645 in the monomeric protein active sites are also labeled. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

essential aspects of the interaction of PADI4 with citrullinatable (or already citrullinated) protein sequences.

In the case of the ARSE peptide (residues Ala204-Lys220), its sequence in intact PADI4 comprises an α-helix (Ser206-Lys210) and a β-strand (Val211-Gln215). Similarly, the GLE peptide (residues Gly375-Arg394) comprises an α-helix (Lys377-Val384) and a β-strand (Gly388-Thr391). In contrast, the PADI-binding regions of other proteins, such as the N-terminal regions of histones [55], are disordered. However, although the elements of secondary structure described above are present in intact PADI4, we have shown in this work that in isolation both the wild-type ARSE and GLE peptides were disordered, and they were still disordered even when single arginine residues were citrullinated (in three cases for the ARSE peptides, and one case for the GLE one). Taking together these findings, we can conclude that citrullination at any of the explored positions of the polypeptide chain does not induce any type of local secondary structure modification, and this conclusion could be only reached by studying the polypeptide regions in isolation, rather than in the secondary and tertiary structure context of the protein. Furthermore, if any structural rearrangement in the citrullinated peptidyl region occurs when the catalytic reaction is carried out in the active site of PADI4, it must solely involve a transient conformational change induced by secondary or tertiary contacts with the catalytic site. It has been shown that auto-deiminated PADI4 (in a one or more yet unidentified arginine positions) modifies its protein-protein interactions, but not its activity [22]. In the specific case of our peptides, the findings we report about the absence of modification of local structure upon citrullination seem also to agree with previous results in which citrullination did not cause any conformational changes, except at the level of quaternary structure or interactions with other proteins [56,57].

Although it is difficult to define the binding location of the peptides, due to their fuzzy nature (i.e., in the absence of a distinct conformation) even in association to PADI4, our simulation results point out that the preferred location for both unmodified arginines or their citrullinated counterparts is along the crevice at the interface of the two protein homodimers of PADI4. We had also previously observed a similar binding to the homodimer interface of PADI4 for short regions of the small intrinsically disordered protein NUPR1 [26]. In fact, it is important to note that this binding region is accessible only to isolated unstructured peptides, as in our case, or larger disordered polypeptide regions. In contrast, due to steric hindrance, the interface crevice is unavailable to structured regions – including these same sequences, either in the unmodified or citrullinated form, when considered in the context of intact PADI4. Therefore, for PADI4 the enzymatic site in each monomer should still be considered the main binding patch for well-folded, intact proteins (especially containing citrullinatable arginines), as well as a potential hot-spot patch for any structured/unstructured molecular partner.

We also measured the kinetic rate constants, k_{on} and k_{off}, for wild-type and citrullinated peptides binding to PADI4. We found that the binding of the peptides in both series was faster (i.e., larger k_{on}) for the unmodified species than for the citrullinated ones. Such a larger value of k_{on} for the wild-type peptides can be directly explained by the presence of the positive charge of the arginine in the wild-type sequence. However, we cannot easily rationalize the fact that the k_{off} values were always slower in the citrullinated peptides than in the wild-type ones.

Although molecular docking is not a technique capable of clarifying directly this point, a possible explanation could be connected to the fact that the binding of either unmodified or citrullinated peptides was observed in simulation at the interface of PADI4 homodimers. In fact, it is known that the three main contributions to the association rate of two polypeptides are the diffusion of the two binding species, long-range electrostatic interactions, and local desolvation effects [58]. In our case, the collision rate determining the diffusion of the binding species can be assumed to be similar between wild-type and citrullinated peptides, and should drive the contact to the most solvent-exposed regions of PADI4 at approximately the same rate. However, due to the different electrostatic charge of citrulline compared to arginine, both Coulombic and desolvation effects play an additional, delicate role in modulating both the association and dissociation. In particular, solvation required to dissociate a peptide from the poorly-hydrated hot-spot binding region of PADI4 observed in simulation (in the crevice of the homodimer interface) could be expected to be slower in the presence of a citrulline residue, due to their lower propensity to coordinate water molecules with respect to arginine. Nevertheless, such a hypothesis needs to be further tested in other cases, which may include the interaction of PADI4 with other substrates (i.e., unmodified or citrullinated regions of other polypeptides).

5. Conclusions

In this study we have investigated the auto-citrullination of PADI4 in an indirect way, by considering peptides encompassing the regions where this protein possesses arginine residues prone to citrullination. In fact, it is not possible to study thermodynamically and kinetically the conformational constraints and consequences of citrullinations by using the whole enzyme due to its large size. The results indicate that the structure of the wild-type peptides, at variance with the conformation found in intact PADI4, is disordered in solution. Furthermore, even when citrullinated, the isolated peptides do not possess any distinct conformation in solution. The binding to PADI4 is predicted to happen at the interface of the homodimeric structure of the protein for citrullinated peptides, although in a heterogeneous (or 'fuzzy') fashion, and takes place in all cases with an affinity in the low micromolar range. However, subtle differences are found in the measured kinetic rate constants, k_{on} and k_{off}, between wild-type and citrullinated peptides in their association to PADI4. Moreover, unmodified peptides may also target the catalytic site, although with a lower affinity. The overall results help to clarify some basic aspects of self-catalysis in PADI4 and, more generally, may aid to elucidate the behavior of other citrullinated proteins, as well as other citrullinating enzymes. Thus, our findings could contribute to the understanding of a key PTM that is increasingly recognized to play a crucial role in many autoimmune and tumoral disorders, since the study of the kinetics and thermodynamics of binding can provide hints on: (i) the design of new inhibitors which could compete for the active site of PADI4 to hamper its auto-citrullination or that of other substrates; and, (ii) the development of new and more efficient ELISAs of auto-antibodies against citrullinated proteins associated with RA patients. This work sightsees the interplay of an enzyme coinciding with its substrate, and therefore it constitutes a step towards understanding some auto-regulation processes connected with citrullination.

CRediT authorship contribution statement

José L. Neira: Writing – review & editing, Writing – original draft, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Bruno Rizzuti:** Writing – review & editing, Writing – original draft, Resources, Methodology, Investigation, Formal analysis, Conceptualization. **Olga Abian:** Writing – review & editing, Writing – original draft, Resources. **Adrian Velazquez-Campoy:** Writing – review & editing, Writing – original draft, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors do not have anything to declare. They do not have any conflict de interest.

Data availability

Data will be made available on request.

Acknowledgements

This research was funded by Spanish Ministry of Science and Innovation MCIN/AEI/10.13039/501100011033/ and "ERDF A way of Making Europe" [PID2021-127296OB-I00 to A.V.C.]; Fondo de Investigaciones Sanitarias from Instituto de Salud Carlos III and European Union (ERDF/ESF, "Investing in your future") [PI21/00394 to O.A.]; by Diputación General de Aragón ["Protein targets and Bioactive Compounds group" E45-23R to AVC, and "Digestive Pathology Group" B25-23R to OA]; and by Consellería de Innovación, Universidades, Ciencia y Sociedad Digital (Generalitat Valenciana) [CIAICO 2021/0135 to JLN]. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

We thank the two anonymous reviewers for helpful suggestions, corrections and discussion. We thank Prof. Ayyalusamy Ramamoorthy and Prof. Kenji Sugase for handling the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bpc.2024.107288.

References

[1] E.R. Vossenaar, A.J.W. Zendman, W.J. Van Venrooij, G.J.M. Pruijn, PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease, Bioessays 25 (2003) 1106–1118.
[2] A.E. Yuzhalin, Citrullination in cancer, Cancer Res. 79 (2019) 1274–1284.
[3] N.S. Gudmann, N.U.B. Hansen, A.C.B. Jensen, M.A. Karsdal, A.S. Siebuhr, Biological relevance of citrullinations: diagnostic, prognostic and therapeutic options, Autoimmunity 48 (2015) 73–79.
[4] A. Ishigami, N. Maruyama, Importance of research on peptidylarginine deiminase and citrullinated proteins in age-related disease, Geriatr. Gerontol. Int. 10 (Suppl) (2010) 1.
[5] B. György, E. Tóth, E. Tarcsa, A. Falus, E.I. Buzás, Citrullination: a posttranslational modification in health and disease, Int. J. Biochem. Cell Biol. 38 (2006) 1662–1677.
[6] J.L. Neira, S. Araujo-Abad, A. Cámara-Artigas, B. Rizzuti, O. Abian, A.M. Giudici, A. Velazquez-Campoy, C. de Juan Romero, Biochemical and biophysical characterization of PADI4 supports its involvement in cancer, Arch. Biochem. Biophys. 717 (2022).
[7] J. Roudier, I. Auger, How does citrullination contribute to RA autoantibody development? Nat. Rev. Rheumatol. 19 (2023) 329–330.
[8] M. Guerrin, A. Ishigami, M.C. Méchin, R. Nachat, S. Valmary, M. Sebbag, M. Simon, T. Senshu, G. Serre, cDNA cloning, gene organization and expression analysis of human peptidylarginine deiminase type I, Biochem. J. 370 (2003) 167–174.
[9] A. Ishigami, T. Ohsawa, H. Asaga, K. Akiyama, M. Kuramoto, N. Maruyama, Human peptidylarginine deiminase type II: molecular cloning, gene organization, and expression in human skin, Arch. Biochem. Biophys. 407 (2002) 25–31.
[10] T. Kanno, A. Kawada, J. Yamanouchi, C. Yosida-Noro, A. Yoshiki, M. Shiraiwa, M. Kusakabe, M. Manabe, T. Tezuka, H. Takahara, Human peptidylarginine

deiminase type III: molecular cloning and nucleotide sequence of the cDNA, properties of the recombinant enzyme, and immunohistochemical localization in human skin, J. Invest. Dermatol. 115 (2000) 813–823.
[11] S. Chavanas, M.C. Méchin, H. Takahara, A. Kawada, R. Nachat, G. Serre, M. Simon, Comparative analysis of the mouse and human peptidylarginine deiminase gene clusters reveals highly conserved non-coding segments and a new human gene, PADI6, Gene 330 (2004) 19–27.
[12] K. Nakashima, T. Hagiwara, A. Ishigami, S. Nagata, H. Asaga, M. Kuramotot, T. Senshu, M. Yamada, Molecular characterization of peptidylarginine deiminase in HL-60 cells induced by retinoic acid and 1α,25-dihydroxyvitamin D3, J. Biol. Chem. 274 (1999) 27786–27792.
[13] S. Dong, T. Kanno, A. Yamaki, T. Kojima, M. Shiraiwa, A. Kawada, M.C. Méchin, S. Chavanas, G. Serre, M. Simon, H. Takahara, NF-Y and Sp1/Sp3 are involved in the transcriptional regulation of the peptidylarginine deiminase type III gene (PADI3) in human keratinocytes, Biochem. J. 397 (2006) 449–459.
[14] D.J. Slade, S. Horibata, S.A. Coonrod, P.R. Thompson, A novel role for protein arginine deiminase 4 in pluripotency: the emerging role of citrullinated histone H1 in cellular programming, Bioessays 36 (2014) 736–740.
[15] E. Witalison, P. Thompson, L. Hofseth, Protein arginine deiminases and associated citrullination: physiological functions and diseases associated with dysregulation, Curr. Drug Targets 16 (2015) 700–710.
[16] S. Ying, S. Dong, A. Kawada, T. Kojima, S. Chavanas, M.C. Méchin, V. Adoue, G. Serre, M. Simon, H. Takahara, Transcriptional regulation of peptidylarginine deiminase expression in human keratinocytes, J. Dermatol. Sci. 53 (2009) 2–9.
[17] Y. Wang, R. Chen, Y. Gan, S. Ying, The roles of PAD2- and PAD4-mediated protein citrullination catalysis in cancers, Int. J. Cancer 148 (2021) 267–276.
[18] C. Yang, Z.Z. Dong, J. Zhang, D. Teng, X. Luo, D. Li, Y. Zhou, Peptidylarginine deiminases 4 as a promising target in drug discovery, Eur. J. Med. Chem. 226 (2021).
[19] M.E. Stensland, S. Pollmann, Ø. Molberg, L.M. Sollid, B. Fleckenstein, Primary sequence, together with other factors, influence peptide deimination by peptidylarginine deiminase-4, Biol. Chem. 390 (2009) 99–107.
[20] F. Andrade, E. Darrah, M. Gucek, R.N. Cole, A. Rosen, X. Zhu, Autocitrullination of human peptidyl arginine deiminase type 4 regulates protein citrullination during cell activation, Arthritis Rheum. 62 (2010) 1630–1640.
[21] M.C. Méchin, F. Coudane, V. Adoue, J. Arnaud, H. Duplan, M. Charveron, A. M. Schmitt, H. Takahara, G. Serre, M. Simon, Deimination is regulated at multiple levels including auto-deimination of peptidylarginine deiminases, Cell. Mol. Life Sci. 67 (2010) 1491–1503.
[22] J.L. Slack, L.E. Jones, M.M. Bhatia, P.R. Thompson, Autodeimination of protein arginine deiminase 4 alters protein-protein interactions but not activity, Biochemistry 50 (2011) 3997–4010.
[23] J. Fuhrmann, K.W. Clancy, P.R. Thompson, Chemical biology of protein arginine modifications in epigenetic regulation, Chem. Rev. 115 (2015) 5413–5461.
[24] J. Fuhrmann, P.R. Thompson, Protein arginine methylation and citrullination in epigenetic regulation, ACS Chem. Biol. 11 (2016) 654–668.
[25] S. Mondal, P.R. Thompson, Protein arginine deiminases (PADs): biochemistry and chemical biology of protein citrullination, Acc. Chem. Res. 52 (2019) 818–832.
[26] S. Araujo-Abad, J.L. Neira, B. Rizzuti, P. García-Morales, C. de Juan Romero, P. Santofimia-Castaño, J. Iovanna, Intrinsically disordered chromatin protein NUPR1 binds to the enzyme PADI4, J. Mol. Biol. 435 (2023).
[27] S. Araujo-Abad, M. Fuentes-Baile, B. Rizzuti, J.F. Bazán, A. Villamarin-Ortiz, M. Saceda, E. Fernández, M. Vidal, O. Abian, A. Velazquez-Campoy, C. de Juan Romero, J.L. Neira, The intrinsically disordered, epigenetic factor RYBP binds to the citrullinating enzyme PADI4 in cancer cells, Int. J. Biol. Macromol. 246 (2023) 125632.
[28] S. Araujo-Abad, B. Rizzuti, A. Villamarin-Ortiz, D. Pantoja-Uceda, C.M. Moreno-Gonzalez, O. Abian, A. Velazquez-Campoy, J.L. Neira, C. de Juan Romero, New insights into cancer: MDM2 binds to the citrullinating enzyme PADI4, Protein Sci. 32 (2023).
[29] S.C. Gill, P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, Anal. Biochem. 182 (1989) 319–326.
[30] J.L. Neira, F. Hornos, J. Bacarizo, A. Cámara-Artigás, J. Gómez, The monomeric species of the regulatory domain of tyrosine hydroxylase has a low conformational stability, Biochemistry 55 (2016).
[31] B. Birdsall, R.W. King, M.R. Wheeler, C.A. Lewis, S.R. Goode, R.B. Dunlap, G.C. K. Roberts, Correction for light absorption in fluorescence studies of protein-ligand interactions, Anal. Biochem. 132 (1983) 353–361.
[32] C.A. Royer, S.F. Scarlata, Fluorescence approaches to quantifying biomolecular interactions, Methods Enzymol. 450 (2008) 79–106.
[33] D. Beckett, Measurement and analysis of equilibrium binding titrations: a beginner's guide, Methods Enzymol. 488 (2011) 1–16.
[34] John Cavanagh, W.J. Fairbrother, A.G.I. Palmer, N.J. Skelton, Protein NMR Spectroscopy: Principles and Practice, Elsevier Science, 1995.
[35] M. Piotto, V. Saudek, V. Sklenář, Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, J. Biomol. NMR 2 (1992) 661–665.
[36] J.L. Neira, F. Hornos, J. Bacarizo, A. Cámara-Artigás, J. Gómez, The monomeric species of the regulatory domain of tyrosine hydroxylase has a low conformational stability, Biochemistry 55 (2016) 3418–3431.
[37] D.K. Wilkins, S.B. Grimshaw, V. Receveur, C.M. Dobson, J.A. Jones, L.J. Smith, Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques, Biochemistry 38 (1999) 16424–16431.
[38] D. M., and W. K, Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of (1)H-(1)H spin-spin coupling constants in proteins. 1983, Biochem. Biophys. Res. Commun. 425 (2012) 519–526.
[39] A. Bax, D.G. Davis, MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy, J. Magnetic Reson. (1969) 65 (1985) 355–360.
[40] A. Kumar, R.R. Ernst, K. Wüthrich, A two-dimensional nuclear overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules, Biochem. Biophys. Res. Commun. 95 (1980) 1–6.
[41] J. Cavanagh, M. Rance, Suppression of cross-relaxation effects in TOCSY spectra via a modified DIPSI-2 mixing sequence, J. Magnetic Resonan. (1969) 96 (1992) 670–678.
[42] M. Kjaergaard, S. Brander, F.M. Poulsen, Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH, J. Biomol. NMR 49 (2011) 139–149.
[43] M. Kjaergaard, F.M. Poulsen, Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution, J. Biomol. NMR 50 (2011) 157–165.
[44] D. Frenzel, D. Willbold, Kinetic titration series with biolayer interferometry, PloS One 9 (2014) e106882.
[45] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2) (2009) 455–461, https://doi.org/10.1002/jcc.21334NA.
[46] B. Rizzuti, W. Lan, P. Santofimia-Castaño, Z. Zhou, A. Velázquez-Campoy, O. Abián, L. Peng, J.L. Neira, Y. Xia, J.L. Iovanna, Design of inhibitors of the intrinsically disordered protein NUPR1: balance between drug affinity and target function, Biomolecules 11 (2021).
[47] B. Rizzuti, J.L. Iovanna, J.L. Neira, Deciphering the binding of the nuclear localization sequence of Myc protein to the nuclear carrier importin α3, Int. J. Mol. Sci. 23 (2022).
[48] J.L. Neira, B. Rizzuti, O. Abián, S. Araujo-Abad, A. Velázquez-Campoy, C. de Juan Romero, Human enzyme PADI4 binds to the nuclear carrier importin α3, Cells 11 (2022).
[49] A. Chakrabartty, T. Kortemme, S. Padmanabhan, R.L. Baldwin, Aromatic side-chain contribution to far-ultraviolet circular dichroism of helical peptides and its effect on measurement of helix propensities, Biochemistry 32 (1993) 5560–5565.
[50] S.M. Kelly, T.J. Jess, N.C. Price, How to study proteins by circular dichroism, Biochim. Biophys. Acta 1751 (2005) 119–139.
[51] S. Vuilleumier, J. Sancho, R. Loewenthal, A.R. Fersht, Circular dichroism studies of barnase and its mutants: characterization of the contribution of aromatic side chains, Biochemistry 32 (1993) 10303–10313.
[52] J. Danielsson, J. Jarvet, P. Damberg, A. Gräslund, Translational diffusion measured by PFG-NMR on full length and fragments of the Alzheimer Aβ(1–40) peptide. Determination of hydrodynamic radii of random coil peptides of varying length, Magn. Reson. Chem. 40 (2002) S89–S97.
[53] J.L. Neira, B. Rizzuti, A. Jiménez-Alesanco, M. Palomino-Schätzlein, O. Abián, A. Velázquez-Campoy, J.L. Iovanna, A phosphorylation-induced switch in the nuclear localization sequence of the intrinsically disordered NUPR1 hampers binding to importin, Biomolecules 10 (2020) 1–22.
[54] G.A. Schellekens, B.A.W. De Jong, F.H.J. Van Den Hoogen, L.B.A. Van De Putte, W. J. Van Venrooij, Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies, J. Clin. Invest. 101 (1998) 273–281.
[55] K. Arita, T. Shimizu, H. Hashimoto, Y. Hidaka, M. Yamada, M. Sato, Structural basis for histone N-terminal recognition by human peptidylarginine deiminase 4, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 5291–5296.
[56] Q. Guo, W. Fast, Citrullination of inhibitor of growth 4 (ING4) by peptidylarginine deminase 4 (PAD4) disrupts the interaction between ING4 and p53, J. Biol. Chem. 286 (2011) 17069–17078.
[57] K. Kizawa, H. Takahara, H. Troxler, P. Kleinert, U. Mochida, C.W. Heizmann, Specific citrullination causes assembly of a globular S100A3 homotetramer: a putative Ca2+ modulator matures human hair cuticle, J. Biol. Chem. 283 (2008) 5004–5013.
[58] C.J. Camacho, S.R. Kimura, C. DeLisi, S. Vajda, Kinetics of desolvation-mediated protein-protein binding, Biophys. J. 78 (2000) 1094–1105.