

RESEARCH

Open Access



Phase separation propensity of the intrinsically disordered AB region of human RXR β

Katarzyna Sołtys^{1*} and Andrzej Ożyhar¹

Abstract

RXR β is one of three subtypes of human retinoid X receptor (RXR), a transcription factor that belongs to the nuclear receptor superfamily. Its expression can be detected in almost all tissues. In contrast to other subtypes – RXR α and RXR γ – RXR β has the longest and unique N-terminal sequence called the AB region, which harbors a ligand-independent activation function. In contrast to the functional properties of this sequence, the molecular properties of the AB region of human RXR β (AB_hRXR β) have not yet been characterized. Here, we present a systematic biochemical and biophysical analysis of recombinant AB_hRXR β , along with *in silico* examinations, which demonstrate that AB_hRXR β exhibits properties of a coil-like intrinsically disordered region. AB_hRXR β possesses a flexible structure that is able to adopt a more ordered conformation under the influence of different environmental factors. Interestingly, AB_hRXR β promotes the formation of liquid–liquid phase separation (LLPS), a phenomenon previously observed for the AB region of another human subtype of RXR – RXR γ (AB_hRXR γ). Although both AB regions seem to be similar in terms of their ability to induce phase separation, they clearly differ in the sensitivity to factors driving and regulating LLPS. This distinct LLPS response to environmental factors driven by the unique amino acid compositions of AB_hRXR β and AB_hRXR γ can be significant for the specific modulation of the transcriptional activation of target genes by different subtypes of RXR.

Keywords Nuclear receptor, Retinoid x receptor subtype beta, AB region, Intrinsically disordered region, Liquid–liquid phase separation, Liquid condensates

Introduction

Retinoid X receptors (RXRs) are members of the nuclear receptor (NR) superfamily of ligand-dependent transcription factors. They were first described in 1992 by Mangelsdorf et al. as NRs able to respond to derivatives of vitamin A [1]. RXRs serve multiple functions in the gene regulation of mammalian signaling systems.

Through the ability to form homodimers, homotetramers and heterodimers with a diverse range of NRs (e.g., peroxisome proliferator-activated receptor (PPAR), liver X receptor (LXR), vitamin D receptor (VDR) and retinoic acid receptor (RAR)), they regulate and coordinate multiple processes, including cell differentiation, development, the immune response, and lipid and glucose metabolism [2]. RXR-controlled transcriptional programs might be potential targets for the treatment of pathologies such as insulin resistance, autoimmunity, atherosclerosis and neurodegeneration [3, 4].

RXRs have characteristic NR architectures that consist of the AB region, a DNA binding domain (DBD), D region, ligand binding domain (LBD) and F region (Fig. 1)

*Correspondence:

Katarzyna Sołtys
katarzyna.soltys@pwr.edu.pl

¹ Department of Biochemistry, Molecular Biology and Biotechnology, Faculty of Chemistry, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

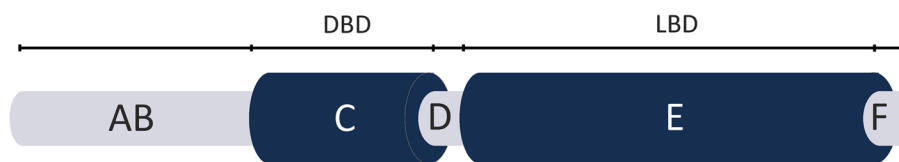


Fig. 1 The structural organization of nuclear receptors. Schematic organisation of nuclear receptors including RXRs. NRs exhibit a modular structure with different regions (A–F). Some of the regions correspond to autonomous functional domains: DNA binding domain (DBD) and ligand binding domain (LBD). Contribution of each domain to the nuclear receptor's activities is different [5]

[5]. NRs share a high degree of sequence homology in their DBDs and LBDs but exhibit weak conservancy in the rest of the protein. Many structures of the DBD and LBD of NRs are currently being determined. However, detailed structural information on the AB region, D region and F region is still lacking. They often exhibit properties of intrinsically disordered regions (IDRs) and are characterized by extraordinary structural flexibility and plasticity [6–10]. IDRs have the ability to undergo fast, highly controllable, environment-modulated transitions that are ideally suited for the transient reversible interactions involved in signal transduction and transcriptional regulation [11, 12]. The N-terminal AB region is involved in the modulation of transcriptional activation of target genes in a cell-specific and promoter-dependent manner. It contains a ligand-independent activation function (AF1) region that is recognized and bound by coactivator proteins or other transcription factors [13]. Flexibility of the D region plays an important structural role by permitting rotation of the DBD and LBD [2]. Little is known about the structure and functional role of the F region. Deletion of the F region may perturb NR activity and interaction with coregulators [14].

There are three RXR subtypes, α , β and γ , which are encoded by three distinct genes [1]. The major variation (both in length and amino acid sequence) among them is in the N-terminal AB region (Fig. S1 and Fig. S2). The three RXR subtypes show tissue-specific expression, with partially overlapping functions [15]. RXR subtype β (RXR β) is widely expressed and can be detected in almost all tissues [1, 16]. RXR β gene transcripts were also detected in a variety of human tumor cell lines [17]. As RXR expression is rarely lost in human tumors, RXR ligands seem to be promising therapeutic targets for cancer therapy and prevention [18]. Loss-of-function studies performed in mice showed that ~50% of RXR β null mutants die in utero [19]. The surviving males were sterile and exhibited testicular defects and abnormal spermatid maturation. The deletion mutants of RXR β may also lead to abnormal lipid metabolism in Sertoli cells [20], which emphasizes the role of functional interactions of RXR β with other NRs that control lipid metabolism.

Polymorphism of the RXR β gene has an effect on hepatitis C virus infection [21] and has been identified in schizophrenia patients [22]. Although detailed functional characterization of RXR β is known, the molecular properties of its AB region, which is critical for several aspects of its action, are still unknown.

An increasing number of reports have shown that some NRs have the ability to liquid–liquid phase separation (LLPS), and the process may be mediated by different regions/domains of NRs. LLPS can be modulated synergistically either by structured domains and/or IDRs of NRs. The structured domains of androgen receptor (AR), PPAR γ and glucocorticoid receptor (GR) are the main drivers of LLPS. In particular, the formation of AR and PPAR γ liquid condensates is mediated by DBD [23, 24], whereas GR is mediated by LBD [25]. On the other hand, for the estrogen receptor (ER) and RXR γ , AB regions seem to be essential for condensate formation [26, 27]. Additionally, LLPS of RXR γ is driven by hydrophobic interactions, which is quite an unusual feature for IDRs [28].

As the major differences in the subtypes of RXR are in their AB regions, the aim of our study was to reveal the molecular properties of the AB region of human RXR β (AB_{hRXRB}). We present in silico examinations with comprehensive biochemical and biophysical characteristics of AB_{hRXRB}. Our data demonstrate that AB_{hRXRB} exhibits some properties of IDRs, but it is also characterized by the presence of ordered secondary structure motifs. In particular, AB_{hRXRB} shows the structural and functional characteristics of the coil-like group of intrinsically disordered proteins (IDPs). Moreover, in the presence of external factors such as osmolyte or increased temperature, AB_{hRXRB} reveals a significant propensity to form additional ordered structures. Most importantly, we demonstrate that AB_{hRXRB}, similar to the AB region of RXR γ (AB_{hRXRG}), exhibits potential for LLPS. The unique amino acid composition of AB_{hRXRB}, especially the high content of P amino acid residues and specific response to different environmental factors driving LLPS, govern the distinct phase separation propensity of AB_{hRXRB}. The differences in

the amino acid sequence and specific response to factors driving LLPS of a particular AB region of RXR can have a substantial impact on the action of individual subtypes of the receptor.

Materials and methods

Chemicals

All buffers were prepared at room temperature. The lysis buffer contained 20 mM Na₂HPO₄ and 150 mM NaCl, pH 7.0. Buffer A contained 30 mM Na₂HPO₄, 500 mM NaCl, 5% (v/v) glycerol, and 0.5 mM DTT, pH 7.0. Buffer B contained 30 mM Na₂HPO₄, 150 mM NaCl, and 5% (v/v) glycerol. Buffer C contained 30 mM Na₂HPO₄, 300 mM NaCl, and 5% (v/v) glycerol, pH 7.0. Buffer D contained 30 mM Na₂HPO₄, 300 mM NaCl, and 1% (v/v) glycerol, pH 7.0. Buffer E contained 20 mM Tris-HCl, 150 mM NaCl, and 5% (v/v) glycerol, pH 7.0. Buffer IEX_A contained 30 mM Na₂HPO₄, 20 mM NaCl, 5% (v/v) glycerol, and 0.5 mM DTT, pH 7.0. Buffer IEX_B contained 30 mM Na₂HPO₄, 1 M NaCl, 5% (v/v) glycerol, and 0.5 mM DTT, pH 7.0. PEGs, TMAO, 1,6-hexanediol, Ficoll 70, trehalose, and phenylmethylsulfonyl fluoride (PMSF) were purchased from Sigma-Aldrich, and (NH₄)₂SO₄ was purchased from Carl Roth.

Preparation of the cDNA construct

The *E. coli* strain DH5 α (Thermo Fisher Scientific) was used as the host strain for the cloning procedures. The sequence of the full-length human RXR β (*hRXR β*) was taken from UniProtKB—P28702. The cDNA of *hRXR β* was de novo synthesized in GeneArt[®] (Thermo Fisher Scientific). The gene sequence was optimized using Gene Optimizer software to maximize the expression of the synthetic gene in *E. coli*. An optimized sequence was used as the template for the PCR. The following primers were used for amplification: forward: CCGGGGCCA TGGGAagctgggcagcagctccgc; reverse: GCGCGCGC GCCGCTTAacctgcaccaggaccaccag. The forward primer introduced the *Nco*I restriction site, and the reverse primer introduced the *Not*I restriction site. The capital letters in the sequences represent nucleotides that are added to coding sequences for cloning purposes, whereas the restriction sites are highlighted. The insert was ligated with the pETHSu plasmid in a frame with the sequence that encodes the N-terminal peptide containing the polyhistidine (6xHis) tag and SUMO [29]. The sequence of the obtained construct was verified by DNA sequencing.

Expression and purification of the AB region of *hRXR β* (AB-*hRXR β*)

AB-*hRXR β* (amino acid residues 1–202) was expressed in the *E. coli* strain ArcticExpress (DE3) (Agilent).

Bacteria harboring the expression vector were grown in ZYM-5052 autoinducing media [30] supplemented with 50 μ g/ml kanamycin. The culture was grown at 37 °C and 182 rpm until the optical density (OD₆₀₀) reached 1.0, and then protein was expressed at 16 °C overnight. The cells were harvested by centrifugation at 5000 \times g at 4 °C for 15 min. The pellet of bacteria was suspended in ice-cold lysis buffer containing PMSF (0.2 mg/ml) and lysozyme (1 mg/ml) and stored at –80 °C until use.

The frozen cell suspension supplemented with DNase (20 μ g/ml), RNase (20 μ g/ml) and a fresh portion of PMSF (0.2 mg/ml) was slowly thawed at 10 °C. Cell lysis was improved by sonication. The cell suspension was sonicated with 5 short 20-s bursts followed by intervals of 30 s for cooling. The resulting suspension was centrifuged at 18 000 \times g at 4 °C for 1 h. The soluble fractions containing AB-*hRXR β* supplemented with PMSF (0.2 mg/ml) were purified using immobilized metal affinity chromatography (IMAC). The cell lysate was incubated for 1 h at 10 °C with Talon[®] Metal Affinity Resin (Clontech), which had been previously equilibrated with lysis buffer. The resin was washed with buffer A and buffer C until the absorbance at 280 nm was lower than 0.1. Then, 0.5 mg SUMO hydrolase (dtUD1) [31] was added to the resin, gently mixed and incubated overnight at 4 °C. The protein was eluted with buffer C, followed by buffer exchange to buffer IEX_A and concentration to a total volume of 500 μ l using Amicon Ultracel-4 Centrifugal Filter Units (Merck Millipore) with a cutoff limit of 10 kDa. AB-*hRXR β* was purified to homogeneity using a MonoS 5/50 GL (GE Healthcare Life Sciences) column pre-equilibrated with buffer IEX_A connected to an ÄKTA Avant system (GE Healthcare Life Sciences). The purification was performed at a 0.5 ml/min flow rate at room temperature. The column was washed with 3 CV of buffer IEX_A followed by a linear gradient of buffer IEX_B for 20 min. Fractions containing the purified AB-*hRXR β* were combined, the buffer was changed to buffer C and the sample was concentrated to 1 mg/ml using Amicon Ultracel-4 Centrifugal Filter Units (Merck Millipore) with a cutoff limit of 10 kDa and then aliquoted into small volumes. The concentration of AB-*hRXR β* was determined spectrophotometrically at 280 nm. The absorption coefficient calculated according to Gill and von Hippel [32] was 0.829 ml/(mg \times cm). The samples were stored at –80 °C. The molecular mass of AB-*hRXR β* was determined in a mass spectrometry laboratory (IBB PAS, Warsaw). All of the experiments presented in this paper were performed using samples obtained from different preparations. The results were reproducible, and we did not observe any variability among the different preparations.

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE)

The protein samples were analyzed by SDS–PAGE using 4%/15% polyacrylamide gels [33]. Electrophoresis was performed at a constant current of 20 mA/1 mm gel with the Unstained Protein Molecular Weight Marker (Thermo Fisher Scientific). After electrophoresis, the gels were stained with Coomassie Brilliant Blue R-250 [34] and analyzed using Image Lab Software (Bio-Rad).

Circular dichroism spectroscopy

Circular dichroism (CD) spectra were recorded using a Jasco-815 spectropolarimeter (Jasco Inc) equipped with a Peltier temperature controller (CDF-426S/15). The spectra were collected in a spectral range of 190–260 nm with a scanning speed of 50 nm/min at 20 °C, D.I. T – 2 s and a 1-nm band width. The data pitch was 0.5 nm, and the final spectrum was obtained after averaging three measurements. The spectra were measured using quartz cuvettes with a path length of 1 mm, and the concentration of AB_hRXRB was 10 μM (0.2 mg/ml) in buffer B. The measurements in the presence of GdmCl, TFE, TMAO, PEG 8000 and (NH₄)₂SO₄ were performed after 30 min of incubation at room temperature. Temperature-dependent denaturation was monitored by following the changes in ellipticity at 222 nm by increasing the temperature from 20 °C to 70 °C and then decreasing it from 70 °C to 20 °C at a constant rate of 1 °C/min. All the spectra were corrected for the effect of the buffer and converted to molar residual ellipticity units. The molar molecular mass per residue (MRW) for AB_hRXRB is 98.99 Da. Evaluation of the secondary structure content was calculated with CDPPro deconvolution software using three algorithms: Continll, Selon3 and Cdsstr. IBasis 6 and 7 (SPD42 and SPD48) were used as the reference protein dataset [35].

Sedimentation velocity analytical ultracentrifugation experiments (SV-AUC)

Sedimentation velocity experiments were performed at 20 °C on a Beckman Coulter ProteomeLab XL-I ultracentrifuge (Beckman Coulter Inc.) in an An-60 Ti rotor. AB_hRXRB (400 μL of protein at three concentrations: 0.25 mg/ml, 0.5 mg/mL, or 1.0 mg/mL in buffer D) was loaded in two-channel centerpieces and centrifuged overnight at 50 000 rpm. Detection of the protein concentration was performed using OD measurements at a wavelength of 280 nm. The data were analyzed with SEDFIT software using a continuous size distribution c(s) model to extract the sedimentation coefficient (s) [36, 37]. The partial specific volumes (V_{bar}) of AB_hRXRB (0.724747 ml/g) and the density (1.0169 g/ml) and

dynamic viscosity (0.010627 mPa×s) of the buffer at 20 °C were calculated using SEDNTERP software [38]. Maximum entropy regularization with $p=0.68$ was applied. The sedimentation coefficient (s), after correction for the solvent density and viscosity in relation to the density and viscosity of water at 20 °C, was expressed as $s_{20,w}$. The hydrodynamic dimensions of AB_hRXRB (R_S and M) were calculated by SEDFIT. The quality of the fits was assessed using the RMSD values, residual distributions and residual histograms.

Limited proteolysis

Purified AB_hRXRB (0.5 mg/ml) was digested with proteinase K (A&A Biotechnology) and trypsin (Lonza) at a final concentration of 25 μg/ml, and the endoproteinase Glu-C (V8) (Sigma–Aldrich) was digested using a 1:250 (w/w) substrate-to-protease ratio in the absence or presence of 30% TFE. The control reactions did not contain any enzymes. The proteolysis reaction was conducted at 15 °C in buffer E. After defined time intervals (5, 15 and 60 min), 10 μl samples were taken, mixed with 3 μl SDS loading buffer and heated for 10 min at 95 °C. Cleaved peptides were resolved using 4%/15% SDS–polyacrylamide gels.

Droplet formation assay

For microscopic and spectroscopic examinations, purified AB_hRXRB was concentrated to 400 μM using Amicon Ultracel-4 Centrifugal Filter Units (Merck Millipore) with a cutoff limit of 10 kDa. The protein was mixed in low-binding microtubes (72.706.600; Sarstedt) with phosphate buffer supplemented with the appropriate amount of sodium chloride, PEGs, TMAO, 1,6-hexanediol, (NH₄)₂SO₄, Fikoll 70 and trehalose. Optical density measurements at 340 nm using a NanoDrop 2000c spectrophotometer (Thermo Scientific) were performed to estimate the turbidity of the protein samples. The droplet formation of the protein samples was monitored using DIC microscopy. Two microliters of the sample was loaded onto glass coverslips, and DIC images were acquired using a Zeiss Axio Observer 7 inverted microscope with a 100× objective (oil immersion). The images were processed using ZEN 3.0 (ZEN lite).

Fluorescent labeling of the protein

The AB_hRXRB protein was fluorescently labeled using the Atto 488 NHS ester (41,698; Sigma–Aldrich) according to the manufacturer's manual. The calculated degree of labelling was 0.7. To confirm the presence of protein in the droplets, Atto 488-labeled protein was mixed with unlabeled protein, and images were acquired on a Zeiss Axio Observer 7 inverted microscope with a 100× objective (oil immersion).

Fluorescence anisotropy measurements

Fluorescence anisotropy measurements were performed on an FP-8300 spectrofluorometer (Jasco) equipped with excitation and emission polarizers and with a Peltier temperature controller (EHC-813). QS High Precision Cell cuvettes (105–251-15–40; Hellma Analytics) were used. All the samples were prepared to a final volume of 60 μ l and were monitored at an excitation wavelength of 498 nm and emission wavelength of 520 nm, with 5-nm bandwidths. The changes in fluorescence anisotropy of the AB_hRXR β samples in buffer B at different pH values were monitored by increasing the temperature from 5 °C to 50 °C at a constant rate of 1 °C/min. The Atto 488-labeled protein was incubated with unlabeled protein (200 μ M) in a 1:100 ratio. Anisotropy (r), including the instrument G factor, was calculated using the Equation $r = \frac{I_{||} - I_{\perp}}{I_{||} + 2 \times I_{\perp}}$, where $I_{||}$ and I_{\perp} are the components of the fluorescence intensity that are parallel and perpendicular to the electric vector of the excitation light, respectively [39].

Fluorescence spectroscopy

8-anilino-1-naphthalenesulfonic acid (ANS) fluorescence measurements were performed on a Fluorolog-3 spectrofluorometer (Horiba Jobin Yvon Inc.) at 20 °C. The concentration of the ANS stock solutions was calculated using $\epsilon_{351} = 6240 \text{ M}^{-1} \times \text{cm}^{-1}$ [40]. The 3 μ M AB_hRXR β samples containing 20 μ M ANS were prepared in buffer E in the presence of 5% PEG 8000 (no LLPS), 0.5 M TMAO (no LLPS), 20% PEG 8000 (LLPS) or 1.5 M TMAO (LLPS). ANS fluorescence was excited at 351 nm, and the emission spectra were recorded in the wavelength range of 400–650 nm. An integration time of 0.5 s and slits with bandwidths of 4 nm were used.

Sequence analysis

Analysis of the AB_hRXR β sequence was performed using bioinformatics tools with the default settings. The ProtParam tool (<https://web.expasy.org/protparam/>) allows for the computation of various physical and chemical parameters (e.g., molecular mass (M), theoretical pI, extinction coefficient) [41]. The amino acid composition was analyzed using a Composition Profiler (<http://www.cprofiler.org>) [42]. The linear net charge per residue of AB_hRXR β was obtained using the CIDER tool (<http://pappulab.wustl.edu/CIDER/analysis/>) [43], and hydrophathy scores were obtained using ExPASy according to the Kyte & Doolittle method (<https://web.expasy.org/protscale/>) [41]. The disordered regions were predicted using IUPred2A (short) (<https://iupred2a.elte.hu/>) [44, 45] and PONDR (VLXT) (<http://www.pondr.com>) [46] predictors. Protein backbone dynamics were

predicted using DynaMine (<http://dynamine.ibsquare.be>) [47, 48]. Charge-hydrophathy analysis and additional disorder prediction were conducted using the PONDR server [46]. To analyze the phase separation propensity of AB_hRXR β , four predictors were used: catGRANULE (http://s.tartagialab.com/update_submission/302811/6360102143) [49], PScore (<http://abragam.med.utoronto.ca/~JFKlab/Software/psp.htm>) [50], PSPredictor (<http://www.pkumdl.cn:8000/PSPredictor/>) [51], and FuzDrop (<https://fuzdrop.bio.unipd.it/predictor>) [52]. The potential sites cleaved by endoproteinase Glu-C (V8), trypsin and proteinase K in the AB_hRXR β sequence were predicted using PeptideCutter (https://web.expasy.org/peptide_cutter/) [41]. Alignment of the three human subtypes of RXR sequences was performed using ClustalX 2.1 [53].

Results

Expression, purification and in silico analysis of the AB region of hRXR β (AB_hRXR β)

To perform biochemical and biophysical analyses of recombinant AB_hRXR β , we elaborated and optimized a protocol for its efficient expression and purification. AB_hRXR β was expressed in the *E. coli* strain Arctic-Express (DE3) from the pETHSu plasmid encoding the cDNA sequence of the AB region of hRXR β with the N-terminal peptide containing the polyhistidine (6xHis) tag and SUMO. Different thermal conditions (16 °C, 20 °C and 30 °C), media (TB and ZYM-5052 autoinducing media) and IPTG induction times (30 min, 1 h, 2 h, 3 h, 5 h) were evaluated to optimize the overexpression of AB_hRXR β (data not shown). After optimization, AB_hRXR β was overexpressed at 16 °C overnight in ZYM-5052 autoinducing medium. A temperature higher than 16 °C resulted in more degradant products of AB_hRXR β (data not shown). To prevent AB_hRXR β degradation, the cell extract was supplemented with PMSE, a serine protease inhibitor. PMSE was added at every working step of the purification procedure until a majority of the contaminant proteins were removed. Additionally, all subsequent steps were carried out on ice or at 4 °C.

Because the pETHSu plasmid enables protein expression with the 6 \times His tag, immobilized metal affinity chromatography (IMAC) using resins containing cobalt ions (Talon[®] Metal Affinity Resin) was chosen as the first step of the purification procedure. Once bound to the resin, AB_hRXR β was washed with a high-salt concentration buffer, which enabled the reduction of contamination without significant protein loss (see Fig. 2B, Lane 5). Next, the ionic strength was lowered, which also enabled the reduction of contaminants (see Fig. 2B, Lane 6) and further digestion with SUMO hydrolase (dtUD1). Preliminary size-exclusion chromatography (SEC) was

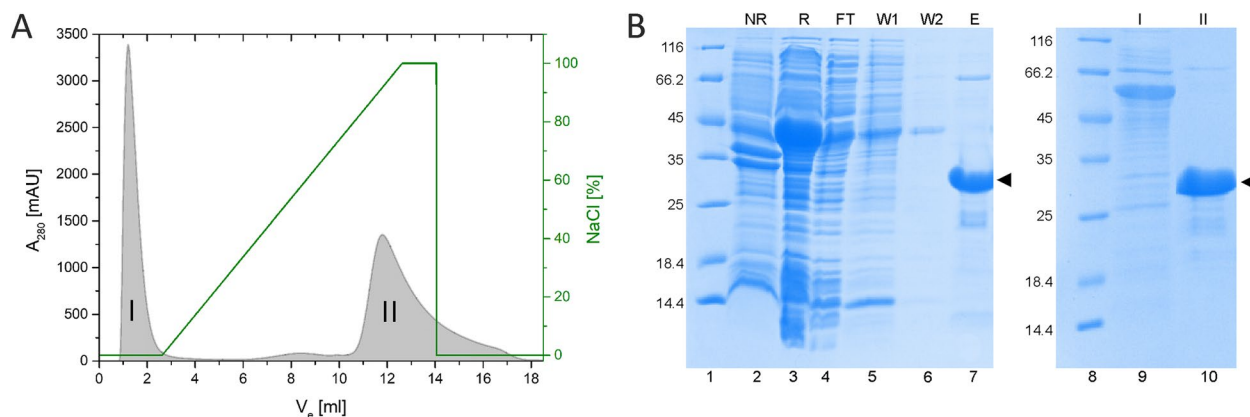


Fig. 2 Purification of AB_{hRXRB}. **A** Ion exchange (IEX) chromatography using a MonoS 5/50 GL column. The absorbance was monitored at 280 nm. The green line corresponds to the gradient of the NaCl concentration. Fractions from the II peak, containing pure AB_{hRXRB}, were combined, concentrated and used for further analysis. **B** SDS-PAGE analysis of the expression and purification of AB_{hRXRB}. The gel was stained with Coomassie Brilliant Blue R250. Lanes 1 and 8, molecular mass standard; Lane 2, the fraction of insoluble proteins obtained after cell lysis; Lane 3, the fraction of soluble proteins obtained after cell lysis; Lane 4, the fraction of proteins unbound to Talon[®] Metal Affinity Resin; Lanes 5 and 6, fraction of proteins eluted with buffers A and C, respectively; Lane 7, the combined fraction after digestion and elution with buffer C; Lane 9, fraction of protein impurities separated during IEX chromatography (peak I); Lane 10, combined fractions of pure AB_{hRXRB} eluted from the MonoS 5/50 GL column (peak II). The arrowhead marks the position of a band corresponding to AB_{hRXRB}

used as an additional step of purification, but unfortunately, it was not possible to obtain a homogenous protein sample (data not shown) using this technique. Thus, ion exchange chromatography with a MonoS column was used (Fig. 2A). This step turned out to be crucial to separate AB_{hRXRB} (Fig. 2A, II peak) from the remaining contaminants, including the degradation products (Fig. 2A, I peak; Fig. 2B, Lanes 9–10). The described purification procedure yielded up to 14 mg of AB_{hRXRB} from 1 L of ZYM-5052 autoinducing culture medium. Purified AB_{hRXRB} appeared as a single band on the SDS-PAGE gel (Fig. 2B, Lane 10). The molecular mass of AB_{hRXRB} was confirmed by electrospray ionization mass spectrometry (ESI-MS) (19 896 Da) and was in agreement with the molecular mass estimated with the ProtParam tool (19 897 Da). Deconvolution of the mass spectrum showed the presence of a second (minor) species (20 060 Da). The mass difference between the peaks is approximately 163 Da, suggesting glycosylation of AB_{hRXRB} [54]. However, additional studies are required to identify and confirm the modification.

Interestingly, SDS-PAGE analysis revealed some abnormal characteristics of AB_{hRXRB}. The molecular mass of AB_{hRXRB} estimated on the basis of mobility in SDS-PAGE was 31.2 kDa. This value is approximately 157% of the theoretical molecular mass value calculated with the ProtParam tool (19.9 kDa) and confirmed in the ESI-MS (19 896 Da). The atypical molecular mobility in the SDS-PAGE experiments is characteristic of intrinsically disordered regions (IDRs), which are known to migrate slower in SDS-PAGE than globular proteins

with the same molecular mass. Because of their unique amino acid composition, IDRs do not bind SDS as much as globular proteins. The apparent molecular mass estimated from SDS-PAGE is often 1.2–1.8 times higher than the real mass calculated on the basis of the primary structure [55].

During the overexpression and purification procedures of AB_{hRXRB}, we observed its high susceptibility to degradation and increased electrophoretic mobility, which might indicate that AB_{hRXRB} exhibits properties of IDRs or intrinsically disordered proteins (IDPs). It has been reported that some AB regions of nuclear receptors (NRs) display characteristic properties of IDRs [6–8, 27]. Intrinsic disorder was suggested to provide benefits to these transcription factors and enable them to carry out their functions. These results inspired us to take a closer look at the molecular properties of AB_{hRXRB}. Analysis of the amino acid composition of AB_{hRXRB} using the Composition Profiler tool [42], which helps to detect the enrichment (values above zero) or depletion (values below zero) of amino acids by their properties, was performed (Fig. 3A). The amino acids are arranged in order of increasing disorder-promoting capacity [56]. The analysis revealed that AB_{hRXRB} is depleted in amino acid residues classified as order-promoting (F and I), and one of them (Y) is completely absent in the sequence. At the same time, the AB_{hRXRB} sequence is rich in amino acid residues characterized as disorder-promoting (G, S, and P). The percentages of P, G and S in AB_{hRXRB} are 25.7%, 15.8% and 11.4%, respectively (Fig. S2). According to the literature, this kind of amino acid distribution

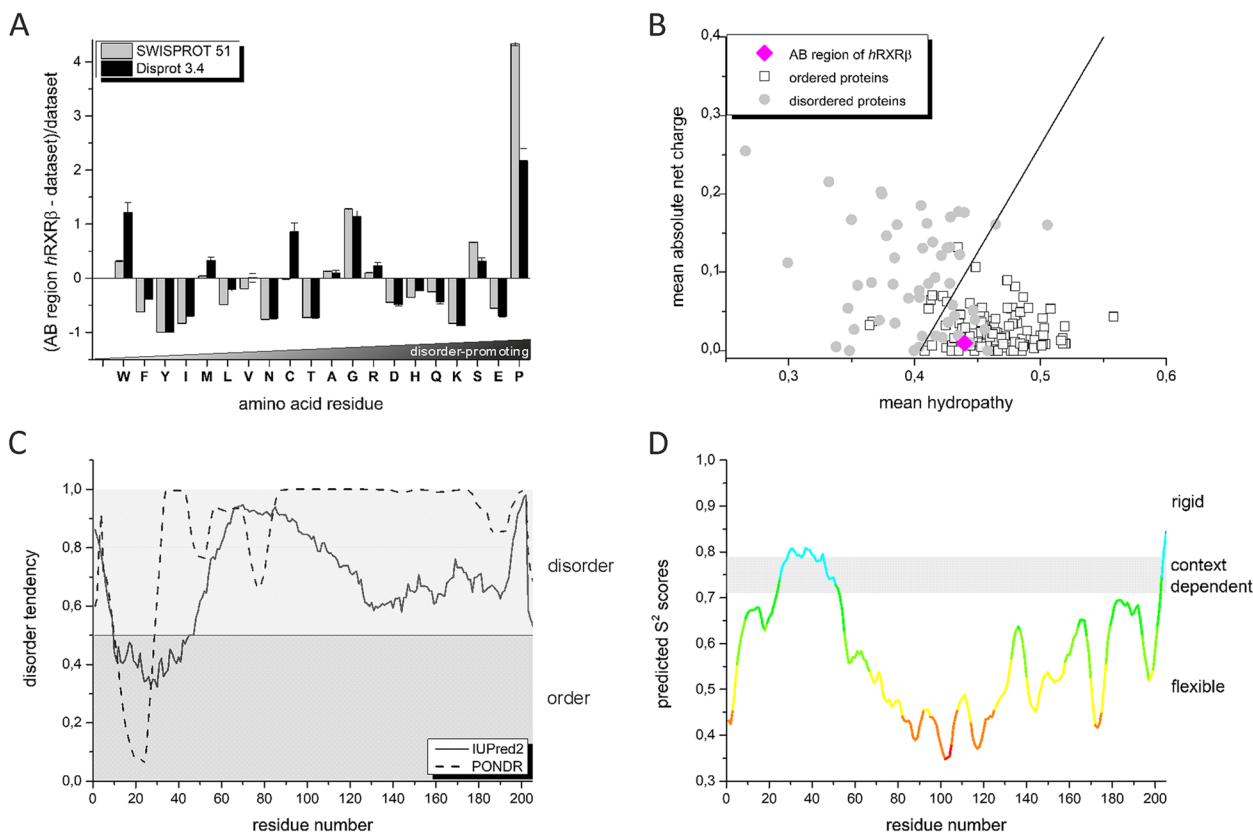


Fig. 3 In silico analysis of the AB_hRXRB sequence. **A** Analysis of the amino acid composition using the Composition Profiler. The plot illustrates the enrichment (values above zero) and depletion (values below zero) of given amino acid residues relative to the proteins from the SwissProt and Disprot 3.4 (dataset with disordered proteins) databases. The amino acid residues are arranged from the most order-promoting to the most disorder-promoting potential. **B** Uversky plot [46] of the mean hydrophathy versus mean absolute net charge of 105 completely ordered proteins (open squares) and 54 completely disordered proteins (gray circles). The solid line represents the border between the ordered and disordered proteins. The magenta diamond corresponds to AB_hRXRB. **C** The prediction of the degree of disorder in the AB_hRXRB sequence calculated from the primary structure. Two algorithms were used – IUPRED2 (short [44, 45] and PONDR (VLXT) [46]. A score above 0.5 indicates a high probability of disorder. **D** Protein backbone dynamics predicted with DynaMine [47, 48]. An S² value larger than 0.8 indicates high rigidity of the protein backbone, whereas an S² value lower than 0.69 indicates high flexibility, which is typical for disordered segments. Values between 0.69 and 0.8 are characteristic of the context-dependent structural organization of polypeptide chains

is typical of IDRs [57]. However, there were some disorder-promoting residues (D, H, Q, K, and E) that were clearly underrepresented in the AB_hRXRB sequence. Additionally, the AB_hRXRB sequence indicated considerable enrichment of W, an order-promoting amino acid residue. In AB_hRXRB, there are three W residues. Thus, analysis of the amino acid composition of AB_hRXRB revealed the dual character of its sequence. In particular, AB_hRXRB possesses a mixture of amino acid residues, which is characteristic of disordered sequences as well as ordered proteins. The charge-hydrophathy plot [58] classified AB_hRXRB into the group of ordered proteins (Fig. 3B). The position occupied by AB_hRXRB is quite ambiguous to interpret, as it emerges within the region occupied by both ordered and disordered proteins.

To estimate the occurrence of the putative intrinsic disorder propensity of AB_hRXRB, in silico analysis using

different disorder predictors was performed. Since the results of most of the predictors were compatible, we decided to show the data obtained for IUPRED2 [44, 45] and PONDR [46] predictors (Fig. 3C). More than 80% of the analyzed sequence seems to be disordered. Only the sequence spanning the 40 amino acid residues in the N-terminus of AB_hRXRB was predicted as a potential ordered region. Additional in silico analysis conducted using DynaMine algorithms (Fig. 3D), a tool that is designated for the prediction of protein backbone dynamics [47], was in agreement with the results obtained using disorder predictors. These data clearly indicate that AB_hRXRB is highly disordered.

Altogether, the unusual properties of AB_hRXRB that were observed, including the differences between the calculated molecular mass value and SDS-PAGE analysis results, the characteristic amino acid sequence

composition and the predicted presence of disordered regions suggest that AB_hRXRB exhibits properties of an IDR.

Hydrodynamic properties of AB_hRXRB

To determine the hydrodynamic properties of AB_hRXRB, sedimentation velocity analytical ultracentrifugation (SV-AUC) experiments were applied. The results are summarized in Table 1. The sedimentation coefficient distributions calculated from the SV-AUC data indicate that one major species with a sedimentation coefficient ($s_{20,w}$) of approximately 1.3 S was observed at three different concentrations (0.25 mg/ml, 0.5 mg/ml and 1 mg/ml) (Fig. 4A); data analysis for all three concentrations yielded a good fit, with RMSD values of 0.007598, 0.009368 and 0.010791, respectively (Table 1 and Fig. S3). No significant dependence of the sedimentation coefficient ($s_{20,w}$) on the protein concentration was observed. The estimated molecular mass (19 144 Da, 21 507 Da and 19 426 Da) was close to the theoretical value (19 897 Da), and the $c(s)$ distribution

Table 1 SV-AUC data obtained for AB_hRXRB

c [mg/ml]	RMSD	f/f_0	M [Da]	R_s [Å]	$s_{20,w}$ [S]
0.25	0.007598	1.955099	19 144	34.5	1.349
0.5	0.009368	2.037723	21 507	37.4	1.398
1.0	0.010791	2.014038	19 426	35.7	1.322

obtained for different protein concentrations overlapped (Fig. 4A), indicating that AB_hRXRB is a monomer. The obtained Stokes radii (R_s) for the different concentrations of AB_hRXRB were 34.5 Å, 37.4 Å and 35.7 Å. These values are considerably higher than the 21.4 Å value calculated from the sequence data assuming that AB_hRXRB is a globular protein (Fig. 4B). The remarkably larger experimental R_s in comparison to the theoretical value places AB_hRXRB on the log (R_s) versus log (M) plot in the area for IDPs with coil-like properties (Fig. 4B). This classification is in agreement with the value of the frictional ratio (f/f_0) obtained in the SV-AUC experiment. Globular proteins are characterized by values of approximately 1.2–1.25, whereas IDPs in which the molecules have extended shapes are characterized by much larger values of f/f_0 [59]. Furthermore, the f/f_0 ratios are typically 2.1 for the 20 kDa and 3.0 for the 200 kDa coil-like IDPs and 1.75 for the 20 kDa and 2.05 for the 200 kDa PMG-like IDPs. The calculated f/f_0 ratio for AB_hRXRB is 2.0 (Table 1), which indicates an extended conformation and classification of AB_hRXRB as a coil-like IDP.

In summary, the data presented above indicate that AB_hRXRB exhibits a highly extended conformation with properties characteristic of coil-like IDPs. Moreover, AB_hRXRB exists as a monomer and does not exhibit a propensity for oligomerization.

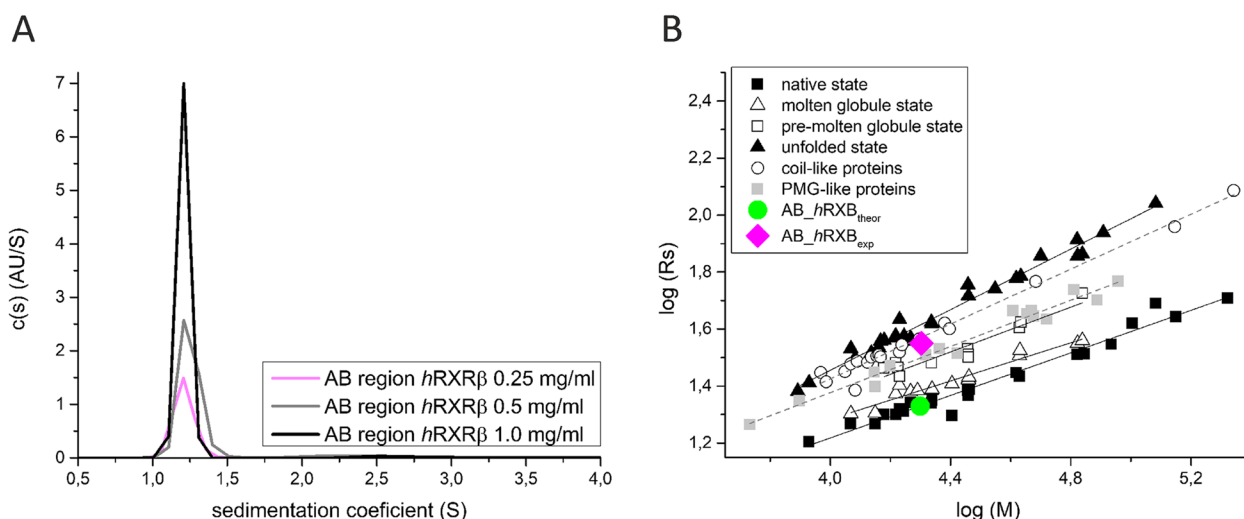


Fig. 4 Hydrodynamic properties of AB_hRXRB. **A** Sedimentation velocity analysis (SV-AUC). Superposition of the sedimentation coefficient $c(s)$ distributions for three different concentrations of AB_hRXRB: 0.25 mg/ml (magenta line), 0.5 mg/ml (gray line) and 1.0 mg/ml (black line). **B** The plot presents the logarithmic values of the hydrodynamic radii (R_s) versus the logarithmic values of the molecular mass (M) for proteins in the native state (black squares), molten globule state (white triangle), premolten globule state (white squares), unfolded state (black triangle), coil-like proteins (white circle) and PMG-like proteins (gray squares). The data for globular proteins were taken from [60], and the data for IDPs are from [58]. The green circle corresponds to the R_s of AB_hRXRB calculated from the equation $\log(R_s^{native}) = (0.373 \pm 0.014) \times \log(M) - 0.274 \pm 0.064$, assuming it is a globular protein (AB_hRXRB_{theor}). The magenta diamond corresponds to the R_s determined for AB_hRXRB after AUC analysis (AB_hRXRB_{exp})

Secondary structure analysis of AB_hRXRB

Circular dichroism (CD) is a method for evaluating the secondary structure and folding properties of proteins. Based on the shape of the CD spectra in far-UV (240–180 nm), it is possible to determine the type of secondary structure present in the analyzed sample [61]. The obtained far-UV CD spectrum of AB_hRXRB is typical for proteins containing disordered regions (Fig. 5A). It is characterized by a deep minimum at 200 nm and a lack of distinct minima at 208 nm and 222 nm. To quantitatively evaluate the content of the secondary structure, the CD spectrum was analyzed by CDPro deconvolution software with IBasis 6 and 7 as the reference protein dataset [35]. Three algorithms were used: Continll, Cdsstr and Selcon3 [35]. Deconvolution of the spectrum revealed that $54.5\% \pm 1.7$ of the sequence is unordered (Table 2). AB_hRXRB also possesses a substantial amount of ordered secondary structure motifs, which could be related to the unique amino acid composition of the sequence (Fig. 3A) and might explain the position of AB_hRXRB on the charge-hydrophathy plot (Fig. 3B). The prevalent type of ordered structure is the β -strand ($18.2\% \pm 2.0$). The

α -helical structure was estimated to be $13.9\% \pm 2.0$ and turns comprise $16.7\% \pm 1.1$ (Table 2).

A double-wavelength plot $[\theta]_{222}$ versus $[\theta]_{200}$ is used to classify IDPs into two structurally different groups: coil-like and PMG-like [62]. Proteins from the first group have an extended conformation that is typical of random coils, whereas the second group consists of proteins that are more compact, although their structure is still less dense than native or molten globule proteins. According to data obtained from the CD spectrum, AB_hRXRB belongs to the group of proteins that are coil-like rather than PMG-like (Fig. 5B).

Data calculated from the CD spectra of AB_hRXRB support the observations obtained from the hydrodynamic property analyses comprising AUC experiments showing that AB_hRXRB exhibits properties of coil-like IDPs.

Folding and unfolding properties of AB_hRXRB

The sensitivity of the protein to differential environmental factors might provide information about its conformation and compactness [63]. Although IDPs/IDRs lack the ability to fold into a stable three-dimensional structure,

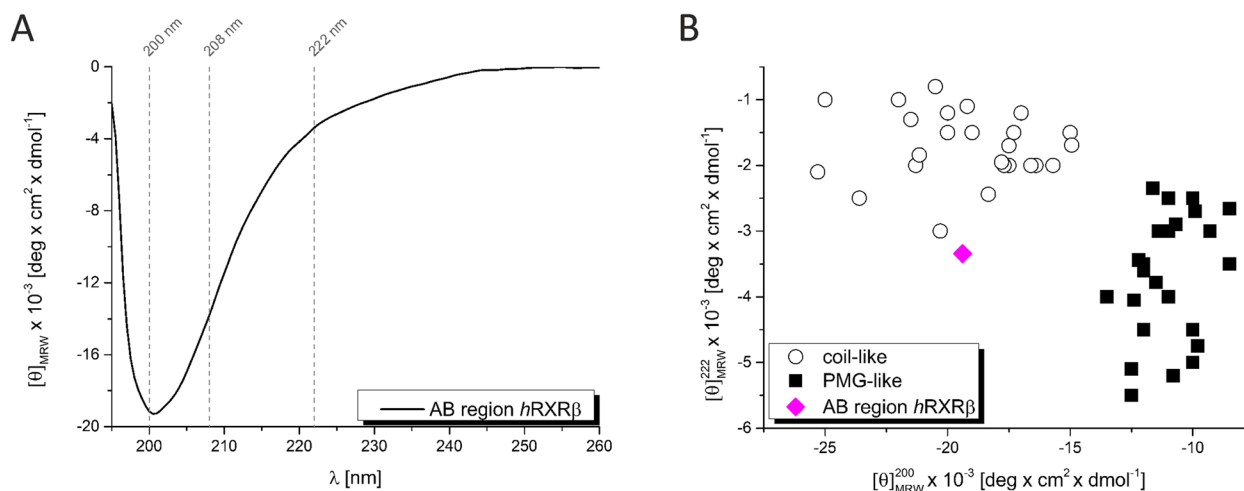


Fig. 5 CD spectrum of AB_hRXRB. **A** The far-UV spectrum of AB_hRXRB. The protein concentration was $10 \mu\text{M}$ (0.2 mg/ml). The spectrum was recorded at 20°C . **B** Classification of the conformational states of AB_hRXRB. The double wavelength plot showing $[\theta]_{222}$ versus $[\theta]_{200}$ of the PMG-like proteins (black squares) and coil-like proteins (circles). The data were taken from [62]. AB_hRXRB is marked with a magenta diamond

Table 2 Estimation of the secondary structure content in AB_hRXRB based on CD spectra

Agent	α -helix (%)			β -strand (%)			Turns (%)	Unordered (%)
	Regular	Distorted	Total	Regular	Distorted	Total		
-	7.5 ± 1.3	6.4 ± 0.7	13.9 ± 2.0	12 ± 1.6	6.2 ± 0.4	18.2 ± 2.0	16.7 ± 1.1	54.5 ± 1.7
30% TFE	10.3 ± 0.9	10.1 ± 1.1	20.4 ± 2.0	7.6 ± 1.3	5.1 ± 0.7	12.7 ± 2.0	17.5 ± 1.3	49.0 ± 3.5
50% TFE	13.1 ± 1.6	17.7 ± 1.4	30.7 ± 3.0	10.9 ± 2.1	9.4 ± 1.2	20.3 ± 3.3	22.5 ± 2.3	23.5 ± 3.1
70°C	10.4 ± 1.6	10.8 ± 0.2	21.2 ± 1.8	14.3 ± 1.8	13.0 ± 2.0	27.2 ± 3.8	23.3 ± 2.2	32.4 ± 3.3

the formation of ordered secondary structural elements, e.g., α -helices, can be induced by different factors [64, 65]. The CD in far-UV in the presence of guanidinium chloride (GdmCl), osmolyte and at various temperatures were recorded to test whether AB_hRXR_B exhibits properties characteristic of IDPs/IDRs and whether it can adopt a more ordered structure in the presence of these factors.

To gain experimental information about the occurrence of the ordered secondary structure, CD measurements were carried out in the presence of the chemical denaturant GdmCl (Fig. 6A). The AB_hRXR_B samples were incubated 1 h before the measurements with an appropriate concentration of GdmCl (0.5 M, 1.0 M and 2.0 M). It was impossible to collect CD data at approximately 200 nm in the presence of GdmCl because of high values of the HT voltage (above 700 V) as well as the quantitative analysis of the whole spectra in the presence of GdmCl with CDPro software. However, changes in the ellipticity at 222 nm with increasing concentrations of GdmCl were observed. When GdmCl concentrations were equal to or below 0.5 M, relatively poor changes in the CD spectrum were observed. However, at higher concentrations of the denaturant (1 M and 2 M), the ellipticity at 222 nm increased, reflecting the loss of the residual ordered structure.

IDPs can form a more ordered structure upon binding to a specific partner (e.g., another protein or a low molecular weight ligand). Osmolytes have become important factors that mimic the in vivo conditions under which IDPs/IDRs interact with target molecules and adopt a more ordered structure [66]. To test the structure-forming potential of AB_hRXR_B, we used 2,2,2-trifluoroethanol (TFE), the osmolyte that contributes to the development of the α -helix structure [67, 68]. Figure 6B shows the impact of increasing TFE concentrations on the far-UV CD spectrum of AB_hRXR_B. At lower concentrations of TFE (5% and 10%), the shapes of the CD spectra were almost the same as those in the absence of osmolytes (Fig. 6B). At 20–25% concentrations of TFE, precipitation of AB_hRXR_B was observed (data not shown). It was reported that TFE can also induce the formation of β -structure-enriched oligomers, which aggregate and form insoluble precipitates [69]. At higher concentrations of TFE (30% and 50%), substantial changes in the CD spectra were observed, which correspond to the induction of an ordered structure accompanied by a loss in the disordered sequence (Fig. 6B and Table 2). The disappearance of the minimum at 200 nm and the formation of two deep ellipticity minima at 208 nm and 222 nm were accompanied by the formation of the secondary structure in AB_hRXR_B. The quantitative analysis of the spectra in the presence of TFE

obtained using CDPro software indicated that the helical content increased from $13.9\% \pm 2.0$ in the absence of TFE to $30.7\% \pm 3.0$ in the presence of 50% TFE (Table 2). The β -strand content first decreased from $18.2\% \pm 2.0$ in the absence of TFE to $12.7\% \pm 2.0$ in the presence of 30% TFE and then increased to $20.3\% \pm 3.3$ in the presence of 50% TFE. The content of the unordered sequence also decreased from 54.5 ± 1.7 in the absence of TFE to 23.5 ± 3.1 in the presence of 50% TFE. The content of turns increased from 16.7 ± 1.1 in the absence of TFE to 22.5 ± 2.3 in the presence of 50% TFE. Thus, AB_hRXR_B underwent a conformational change in the presence of TFE, giving rise to an ordered structure with a prevalence of α -helix.

The analysis of the temperature effect on the structural properties of IDPs reveals that they exhibit a so-called turned out response to heat, where increasing temperatures induce the formation of a secondary structure [63]. AB_hRXR_B displays such properties, as illustrated in Fig. 6C and D, which show the temperature-induced changes in the far-UV CD spectrum of AB_hRXR_B and the temperature-dependence of the ellipticity at 222 nm (Fig. 6E). As discussed above, at low temperatures, the CD spectrum of AB_hRXR_B is typical of an unfolded polypeptide chain. As the temperature was increased, the shape of the spectrum changed, reflecting the temperature-induced formation of ordered secondary structure motifs (Fig. 6C). Additionally, as the temperature increased (from 20 °C to 70 °C), the values of $[\theta]_{222}$ gradually decreased (Fig. 6E; magenta triangles). The changes in the $[\theta]_{222}$ versus temperature plot are linear, suggesting noncooperative folding. The structural heating-induced changes in AB_hRXR_B were completely reversible (Fig. 6E; blue triangles). The CD spectrum recorded at 20 °C for AB_hRXR_B that was previously heated to 70 °C (Fig. 6C) and then cooled to 20 °C (Fig. 6D) perfectly matched the CD spectrum of the sample before denaturation (Fig. 6C). Quantitative analysis of the denatured spectrum at 70 °C was performed with CDPro software (Table 2). Compared to data obtained for AB_hRXR_B at 20 °C, the results show substantial changes in the content of α -helix (from $13.9\% \pm 2.0$ at 20 °C to $21.2\% \pm 1.8$ at 70 °C) and unordered sequence (from $54.5\% \pm 1.7$ at 20 °C to $32.4\% \pm 3.3$ at 70 °C). Moreover, changes were also observed in the content of β -strands (from $18.2\% \pm 2.0$ at 20 °C to $27.2\% \pm 3.8$ at 70 °C) and turns (from $16.7\% \pm 1.1$ at 20 °C to $23.3\% \pm 2.2$ at 70 °C). The analyses presented above clearly demonstrate that AB_hRXR_B exhibits properties characteristic of IDPs/IDRs in response to increased temperature.

Altogether, the above results indicate that AB_hRXR_B is an IDR that has a flexible and dynamic structure. The AB_hRXR_B structure is very sensitive to external factors

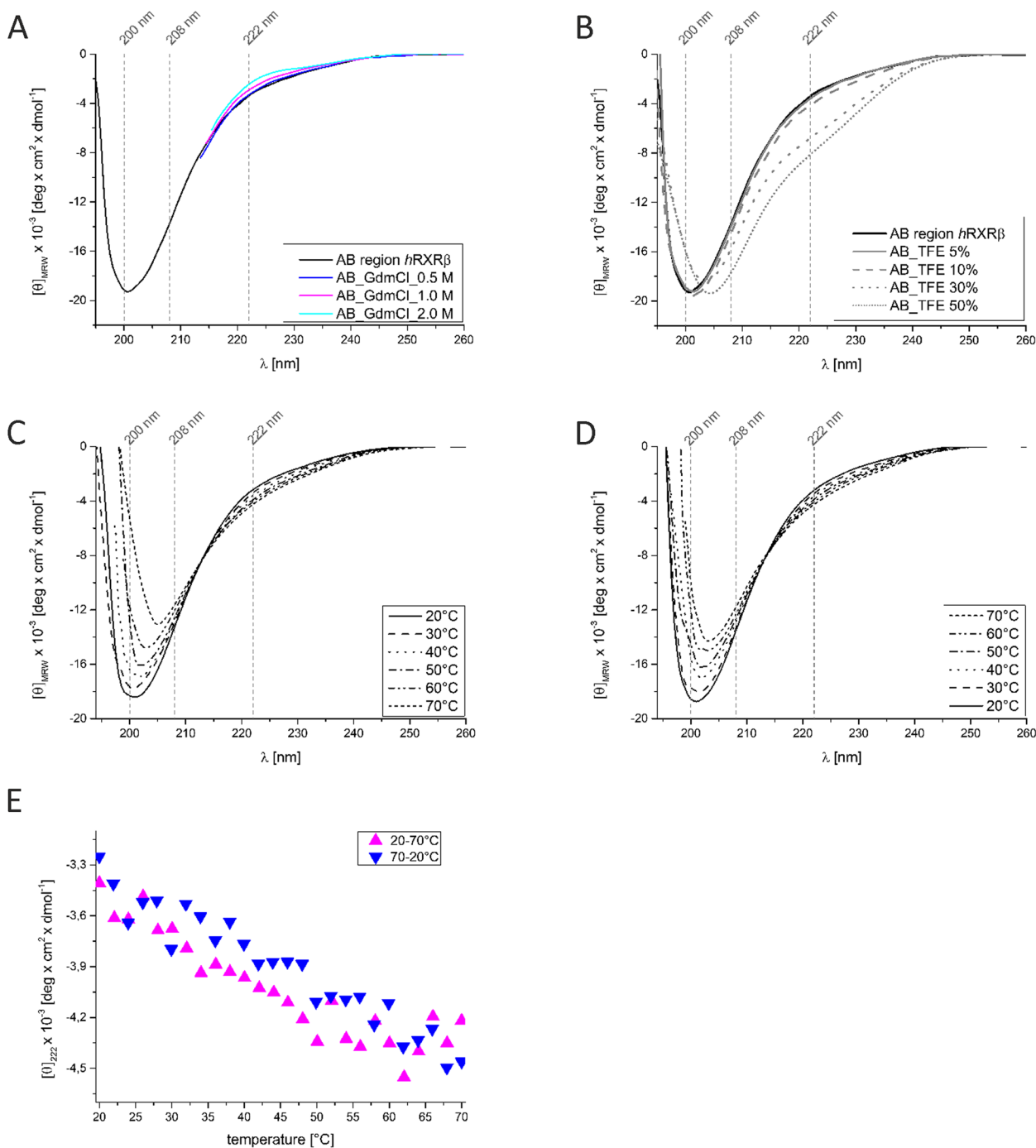


Fig. 6 The effects of GdmCl, TFE and temperature on the content of the secondary structure of AB_{hRXRB}. **A** Far-UV CD spectra of AB_{hRXRB} in the absence (black) and presence of increasing concentrations of GdmCl (0.5 M – blue, 1.0 M – magenta and 2.0 M – cyan). **B** Far-UV CD spectra of AB_{hRXRB} in the absence (black) and presence of increasing concentrations of TFE. **C** Temperature-induced secondary structure formation in AB_{hRXRB}. Representative CD spectra of AB_{hRXRB} measured at temperatures from 20 °C to 70 °C and **D**, from 70 °C to 20 °C. **E** Temperature-induced changes in the CD spectrum ($[\theta]_{222}$ of AB_{hRXRB} versus temperature) measured at temperatures from 20 °C to 70 °C (magenta triangles) and from 70 °C to 20 °C (blue triangles) at 1 °C intervals

and can be easily modulated by them, leading to the formation of ordered secondary structure motifs.

Effect of TFE on the digestion resistance of AB_hRXRB

Flexible protein regions are known to be easily accessible targets for proteases. The addition of binding partners or osmolytes often leads to a protein conformation that is less solvent-exposed and thus less sensitive to protease degradation [70]. To examine changes in AB_hRXRB conformation, three proteases (proteinase K, trypsin and endoproteinase Glu-C (V8)) were chosen to test the proteolytic cleavage of AB_hRXRB in the absence and presence of 30% TFE. Moreover, we wanted to verify whether and to what extent the putative cleavage sites are accessible to proteases. The concentration of the osmolyte was chosen based on the CD data (Fig. 6B) to select the concentration at which substantial changes in the CD spectrum were observed compared to the lower concentration of TFE. AB_hRXRB possesses 56 putative cleavage sites for proteinase K, 9 for trypsin, and 6 for V8 (Fig. 7). The proteolytic susceptibility of AB_hRXRB to proteinase K,

trypsin and V8 in the absence and presence of 30% TFE was analyzed by SDS-PAGE (Fig. 7). AB_hRXRB alone or in the presence of 30% TFE was stable during 1 h of incubation (Fig. 7, Lanes 1–2). AB_hRXRB in the absence of osmolyte appeared to be very sensitive to proteinase K digestion (Fig. 7; Lanes 4, 11, 17), trypsin (Fig. 7; Lanes 6, 13 and 19) and V8 (Fig. 7; Lanes 8, 15 and 21). Moreover, we observed the difference in susceptibility of AB_hRXRB sequence to chosen proteases in the absence of TFE. Although AB_hRXRB possesses 56 putative cleavage sites for proteinase K, we still observed cleavage products with a relatively high molecular mass (Fig. 7; Lane 11), comparing to reaction in the presence of trypsin (9 putative cleavage sites) where after 15 min the digestion products were not detected (Fig. 7; Lane 13). On the other hand, in the presence of 30% TFE we observed protection against proteinase K (Fig. 7; Lanes 5, 12 and 18) trypsin (Fig. 7; Lanes 7, 14 and 20) and V8 (Fig. 7; Lanes 9, 16 and 22) action when compared to AB_hRXRB in the absence of osmolyte. The results showed that in the presence of TFE, AB_hRXRB sequence was less solvent-exposed and

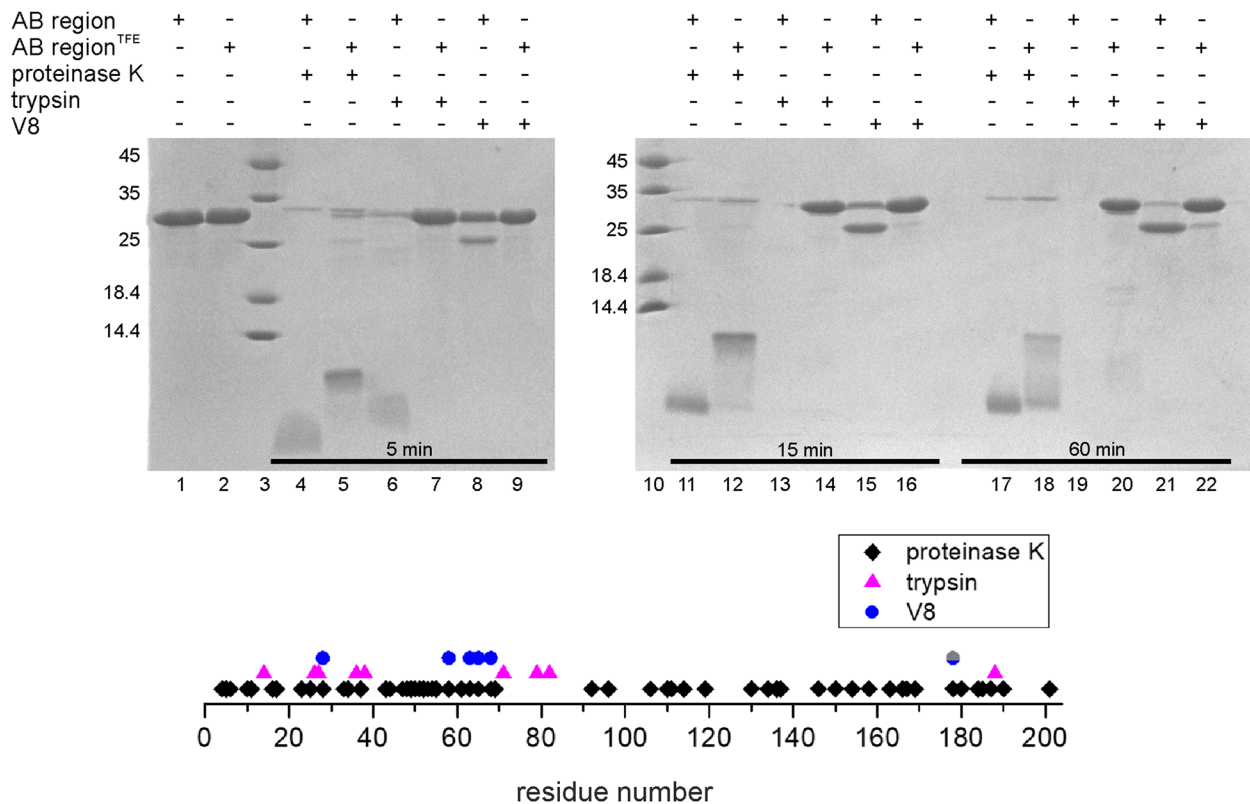


Fig. 7 Effects of TFE on the proteolysis of AB_hRXRB by proteinase K, trypsin and endoproteinase Glu-C (V8). Purified AB_hRXRB alone or in the presence of 30% TFE was digested with proteinase K (Lanes 4–5; 11–12 and 17–18), trypsin (Lanes 6–7; 13–14 and 19–20), or endoproteinase Glu-C (V8) (Lanes 8–9; 15–16 and 21–22). Proteolysis was conducted at 15 °C for an hour. The control samples were the intact AB_hRXRB without TFE (Lane 1) and that in the presence of 30% TFE (Lane 2). The molecular weight marker is shown on the left in kDa (Lanes 3 and 10). Aliquots were taken at 5 min (Lanes 4–9), 15 min (Lanes 11–16) and after 60 min (Lanes 17–22) after the addition of enzymes. The scheme in the lower panel depicts the putative protease cleavage sites of AB_hRXRB

thus less sensitive to protease degradation. Better protection was observed for trypsin and V8, as the substrate sequence possesses only 9 and 6 potential cleavage sites, respectively. This restricted digestion of AB_ *hRXRB* in the presence of 30% TFE along with the CD data confirm that AB_ *hRXRB* adopts a folded conformation in the presence of this osmolyte.

Phase separation propensity of AB_ *hRXRB*

The major differences in the sequence of the subtypes of *hRXR* are located in the AB region (Fig. S1 and Fig. S2). The AB region of NRs contributes significantly to the cell and tissue specificity of the action of NRs. It was previously shown that the AB region of *hRXR γ* (AB_ *hRXRG*) has a propensity for liquid–liquid phase separation (LLPS) [27]. Thus, we decided to investigate whether AB_ *hRXRB* has phase separation propensity, as it might be an important element defining the differences in *hRXR* action. The sequence of AB_ *hRXRB* was subjected to three LLPS predictors: catGRANULE [49], PScore [50], and FuzDrop [52] (Fig. 8). Each of these algorithms uses different criteria (e.g., sequence composition, disorder propensities, nucleic acid-binding properties, propensity for long-range planar π - π contacts) and gives score values indicating the protein's propensity to LLPS. The catGRANULE, PScore and FuzDrop analyses indicated two (54–97 and 150–202 amino acid residues), one (79–99

amino acid residues), and three (1–17, 53–178 and 194–202 amino acid residues) fragments of the AB_ *hRXRB* sequence with positive residue score values, respectively. Additionally, each of these tools gave positive total propensity scores to AB_ *hRXRB* (Table 3). The FuzDrop analysis also showed the presence of probable droplet-promoting regions (data not shown) in the AB_ *hRXRB* sequence; along with the high total propensity score (values above 0.6), this led to the classification of AB_ *hRXRB* as a droplet-driver protein. In summary, analyses using different predictors clearly demonstrate that AB_ *hRXRB* has a propensity to induce LLPS.

Different types of transient interactions may underlie and stabilize liquid condensates, including electrostatic, dipole–dipole, π - π , cation- π , hydrophobic,

Table 3 Total propensity scores for the LLPS of the AB_ *hRXRB* sequence

Predictor	Total propensity score
catGRANULE [49]	0.914
PScore [50]	4.30
FuzDrop [52]	0.9984
PSPredictor [51]	0.9913

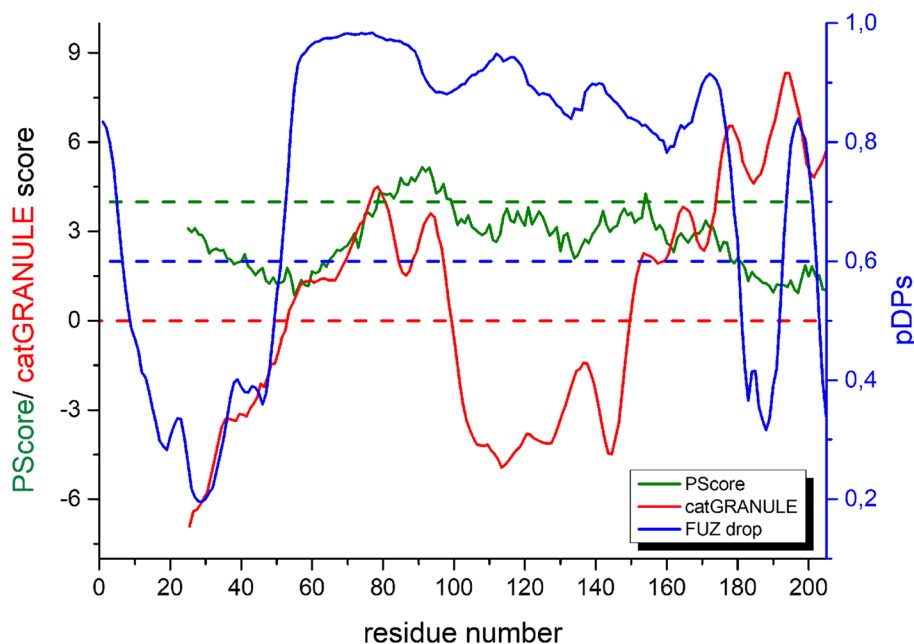


Fig. 8 Prediction of the phase separation propensity of the AB_ *hRXRB* sequence. Graph illustrating catGRANULE (red line), PScore (green line) and FuzDrop (blue line) analysis results for the AB_ *hRXRB* sequence. Values above 0 (red dashed line) and above 4 (green dashed lines) indicate a putative tendency of protein fragments for driving LLPS predicted by the catGRANULE and PScore predictors, respectively, whereas values (pDPS – droplet-promoting probabilities) above 0.6 (blue dashed line) indicate a probable tendency predicted by FuzDrop

and hydrogen bonding interactions [71, 72]. The degree to which side chain and backbone interactions contribute to LLPS depends on the amino acid composition and the overall sequence patterns of the protein [73]. We asked whether the primary structure of AB_hRXRB might indicate a mechanism for the formation of liquid condensates. The AB_hRXRB sequence is abundant in P (25.7%), G (15.8%) and S (11.4%) amino acid residues (Fig. S2). It also contains a motif consisting of seven A amino acid residues. Using the Classification of Intrinsically Disordered Ensemble Regions (CIDER) tool [73], we obtained a linear net charge per residue (NCPR) plot (Fig. 9A), which indicates fractions of negatively and positively charged residues located mainly in the N-terminus of AB_hRXRB, followed by a hydrophobic patch. The calculated values place AB_hRXRB on the Das-Pappas phase diagram in the group of weak polyampholytes (Fig. S4). A Kyte & Doolittle plot for AB_hRXRB, which represents the average hydrophathy along the amino acid sequence, showed that the sequence is hydrophilic rather than hydrophobic (Fig. 9B). The AB_hRXRB sequence has a high content of P amino acid residues, for which the individual value on the hydrophathy/Kyte & Doolittle scale is -1.600 . Despite the hydrophobic side chain, the P amino acid residue is very hydrophilic. Altogether, the above results indicate that the first-order structure of AB_hRXRB is a combination of amino acid residues, which can indicate that a variety of molecular interactions might contribute to the formation of protein liquid condensates.

As *in silico* analysis of the AB_hRXRB sequence did not provide clear information about the types of interactions

driving LLPS, we experimentally determined the phase separation propensity of AB_hRXRB. First, to probe the contribution of electrostatic interactions, we analyzed a solution containing AB_hRXRB at a wide range of protein concentrations and ionic strengths (NaCl concentrations) (data not shown) utilizing both turbidity measurements (optical density at 340 nm) and differential interference contrast (DIC) microscopy. However, AB_hRXRB did not undergo LLPS in response to changing NaCl concentration. This result suggests that electrostatic interactions do not contribute to the LLPS of AB_hRXRB.

We also systematically tested the temperature dependence of the putative AB_hRXRB phase separation by measuring the fluorescence anisotropy of Atto-488-labeled AB_hRXRB (Fig. 10). Fluorescence anisotropy is related to rotational flexibility and increases during the formation of supramolecular assemblies because of the restricted mobility within the assemblies [74]. We tested AB_hRXRB at three different pH values (Fig. 10). However, none of these conditions allowed for us to observe LLPS formation.

Finally, we decided to test the propensity of AB_hRXRB for the formation of liquid condensates in conditions that mimic the crowded intracellular environment. Analysis in a buffer supplemented with various macromolecular crowding agents was performed (Fig. 11A-C). First, we compared the propensity of AB_hRXRB for LLPS in solutions containing PEGs of different molecular masses (Fig. 11A). When AB_hRXRB was added to buffers containing 10% (w/v) PEG 8000 and 10% (w/v) PEG 6000, the formation of condensates by AB_hRXRB was observed and confirmed by turbidity measurements (data

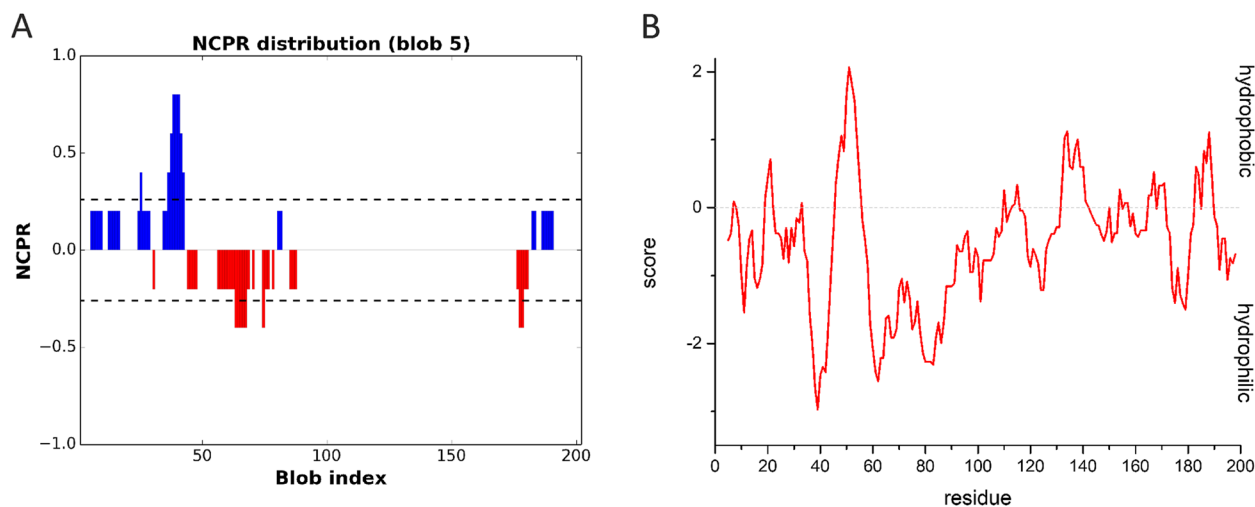


Fig. 9 Primary sequence analysis of AB_hRXRB. **A** Linear net charge per residue (NCPR) plot of AB_hRXRB using the CIDER tool [43]. Positive net charges (blue) and negative net charges (red) are represented. **B** The Kyte & Doolittle plot [41] estimates the hydrophathy scores of the AB_hRXRB residues. Residues with positive scores are predicted to be hydrophobic, while residues with negative scores are predicted to be hydrophilic

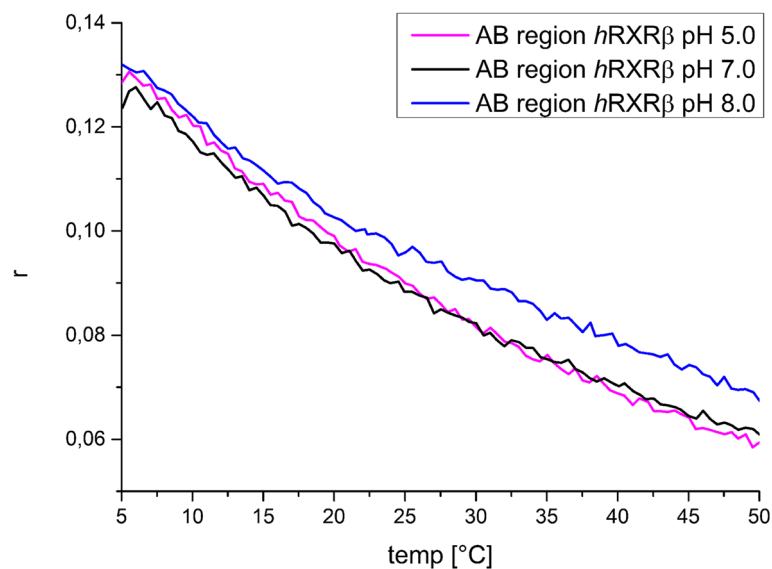


Fig. 10 The effects of temperature and pH on the fluorescence anisotropy of AB_ *hRXRβ*. Temperature-induced changes in the fluorescence anisotropy (r) of Atto 488-labeled AB_ *hRXRβ* measured at temperatures of 5 °C to 50 °C. Measurements were performed at pH 5 (red line), pH 7 (black line) and pH 8 (blue line). The protein concentration was 200 μ M

not shown) and DIC (Fig. 11A). Additionally, differences in the number and size of the droplets were observed (Fig. 11A). In the presence of 10% (w/v) PEG 8000, there were more AB_ *hRXRβ* droplets with sizes between 2–5 μ m, and fewer in the presence of 10% (w/v) PEG 6000, and their size was reduced (1–2.5 μ m). In the case of 10% (w/v) PEG 3000, we did not observe AB_ *hRXRβ* LLPS (Fig. 11A). These data show that the LLPS of AB_ *hRXRβ* is dependent on the molecular mass of the crowding agent. Moreover, the formation of condensates by AB_ *hRXRβ* in the presence of PEGs was observed after 5–10 min, depending on the concentrations of the protein and crowding agent (data not shown). The analysis of the AB_ *hRXRβ* sequence with the PSPredictor tool [51], which gives access to the most similar proteins (to the query protein) in the liquid–liquid phase separation database (LLPSDB) and their experimentally validated phase separation conditions, indicated that for most of the proteins (e.g., Mid1p, Shugoshin (Sgo), Intersectin-1 and son of sevenless homolog 1 (Sos1)), time was a crucial factor in the formation of condensates. The LLPS of AB_ *hRXRβ* was dependent on the molecular mass of PEGs as well as on the crowding agent used. We did not observe droplet formation in the presence of 10% (w/v) Ficoll 70 [75] (Fig. 11B). The formation of liquid AB_ *hRXRβ* condensates was also tested in the presence of 10% (w/v) trehalose (Fig. 11B). However, under these conditions, we were not able to detect droplet formation. We also verified AB_ *hRXRβ* LLPS in the presence of an osmolyte, trimethylamine N-oxide (TMAO) (Fig. 11C).

After the addition of AB_ *hRXRβ* to buffer containing TMAO, the solution immediately turned opaque. LLPS of AB_ *hRXRβ* was confirmed by turbidity measurements (data not shown) and DIC (Fig. 11C). Additionally, using Atto 488-labeled protein, the presence of AB_ *hRXRβ* in the liquid droplets was confirmed. Altogether, the above results indicate that various crowding agents can have different impacts on the formation of AB_ *hRXRβ* condensates.

After confirmation of the propensity of AB_ *hRXRβ* for droplet formation, we decided to ascertain the type of interactions that are responsible for their formation. According to the Hofmeister series of salts [76], chaotropic salts such as GdmCl can weaken the stability of protein assemblies that are stabilized by hydrophobic interactions, whereas “neutral” salts, such as NaCl, or kosmotropic salts such as ammonium sulfate $(\text{NH}_4)_2\text{SO}_4$ can promote hydrophobic interactions between proteins [77]. The formation of AB_ *hRXRβ* liquid condensates was prevented when chaotropic salts such as GdmHCl were applied (data not shown). On the other hand, when a kosmotropic salt such as $(\text{NH}_4)_2\text{SO}_4$ was added, AB_ *hRXRβ* liquid condensates appeared (Fig. 11D), which was confirmed by turbidity measurements and DIC. Additionally, after a few minutes, droplets settled onto and wetted the surface of the glass coverslip, where they remained immobile (Fig. 11D). To obtain further insight into the type of interactions that are responsible for AB_ *hRXRβ* LLPS, we investigated the sensitivity to 1,6-hexanediol, a compound known to disrupt liquid-like condensates

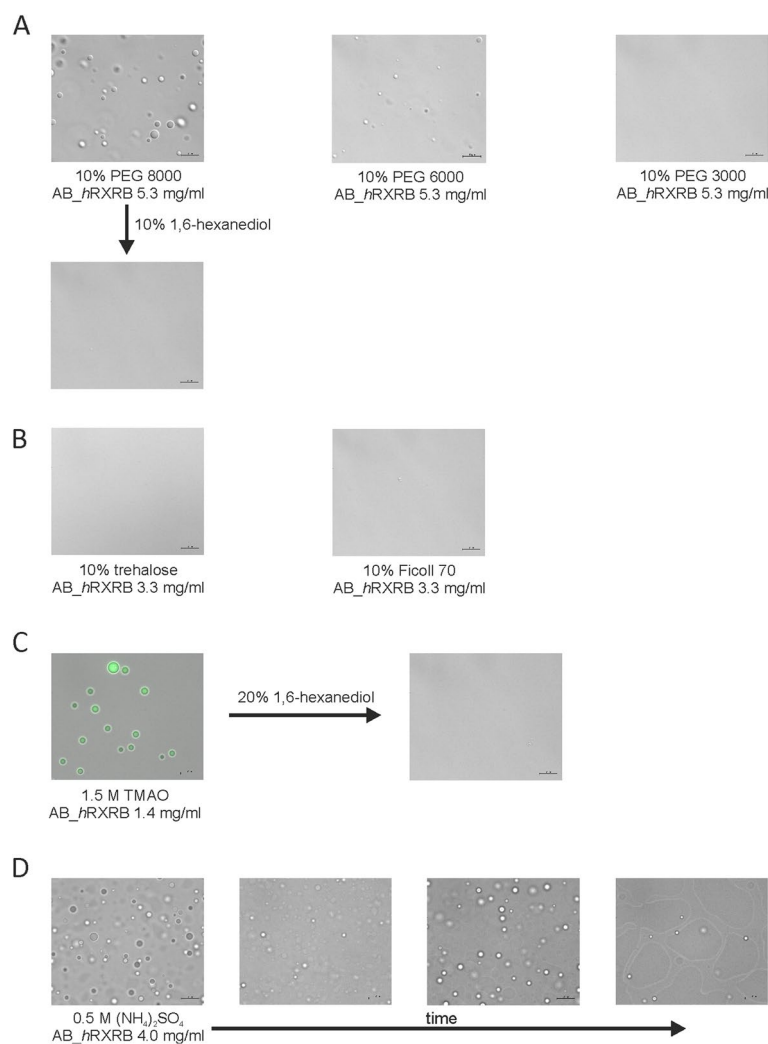


Fig. 11 Factors influencing the formation of AB_hRXRB liquid condensates. **A** The phase behavior of AB_hRXRB in the presence of crowding agents. Representative images obtained by DIC of 260 μ M AB_hRXRB in the presence of 10% (w/v) PEG 8000, 10% (w/v) PEG 6000 and 10% (w/v) PEG 3000. The addition of 10% (w/v) 1,6-hexanediol to the AB_hRXRB sample with 10% (w/v) PEG 8000 allowed for the droplets to dissolve. **B** The phase behavior of AB_hRXRB in the presence of 10% trehalose and 10% Ficoll 70. Representative images obtained by DIC of 165 μ M AB_hRXRB. **C** AB_hRXRB droplet formation in the presence of TMAO. Incorporation of Atto 488-labeled AB_hRXRB (70 μ M) into the droplets in the presence of 1.5 M TMAO. The addition of 20% (w/v) 1,6-hexanediol to the AB_hRXRB sample with 1.5 M TMAO dissolved the droplets. **D** The phase behavior of AB_hRXRB in the presence of kosmotropic salt. LLPS driven by 200 μ M AB_hRXRB in the presence of 0.5 M (NH₄)₂SO₄. The wetting properties of AB_hRXRB liquid droplets on the surface of a glass slide over time are also apparent. Scale bars: 10 μ m

by interfering with hydrophobic interactions [78]. We tested the effects of 1,6-hexanediol on AB_hRXRB liquid condensates in the presence of PEG 8000 (Fig. 11A) and TMAO (Fig. 11C). The turbidity measurements (data not shown) and DIC show that AB_hRXRB liquid condensates were dissolved completely upon the addition of 10% or 20% 1,6-hexanediol to the solution containing PEG 8000 and TMAO, respectively. Lower concentrations of 1,6-hexanediol only partially disrupted the AB_hRXRB liquid condensates (data not shown). The sensitivity of the AB_hRXRB liquid condensates to the presence of

1,6-hexanediol and kosmotropic salt demonstrates that hydrophobic interactions contribute to the LLPS.

We also verified whether the formation of liquid AB_hRXRB condensates in the presence of (NH₄)₂SO₄, TMAO and PEG 8000 is connected with conformational changes in AB_hRXRB. The content of the secondary structures of AB_hRXRB in the presence of these factors was investigated using far-UV CD (Fig. S5). The obtained CD spectra of AB_hRXRB in the presence of (NH₄)₂SO₄ and PEG 8000 were identical for AB_hRXRB (alone) and were typical for proteins containing disordered

regions. These CD data suggest that the protein remains largely disordered within liquid condensates induced by $(\text{NH}_4)_2\text{SO}_4$ and PEG 8000. However, the obtained CD spectrum for AB_hRXRB in the presence of TMAO (Fig. S5) is different from those in the presence of $(\text{NH}_4)_2\text{SO}_4$ and PEG 8000. Changes in the secondary structure in the presence of TMAO and the potential impact of hydrophobic interactions in droplet formation of AB_hRXRB prompted us to look for further structural changes and exposure of hydrophobic amino acid residues. We utilized 8-anilino-1-naphthalenesulfonic acid (ANS) fluorescence in the presence of both PEG and TMAO (Fig. 12) and investigated two concentrations for each factor: one driving AB_hRXRB LLPS (P) and the other not driving AB_hRXRB LLPS (N). Free ANS excited at 351 nm had a fluorescence spectrum characterized by a low emission maximum located at 525 nm. However, the ANS quantum yield increases and the emission maximum shifts toward the blue as the solvent polarity decreases or upon binding to hydrophobic surfaces of proteins [79]. As shown in Fig. 12A, the fluorescence intensity of ANS increased in the presence of 5% and 20% PEG 8000 regardless of the addition of AB_hRXRB. Stronger fluorescent signals were observed for higher PEG 8000 concentrations. Additionally, the position of the fluorescence maxima blueshifted from 525 to 501 nm. However, the changes in the spectra were due to less solvent polarity.

Analogous experiments were conducted in the presence of TMAO. As shown in Fig. 12B, upon addition of AB_hRXRB in the presence of 0.5 M TMAO, the ANS fluorescence intensity increased 1.5-fold, and the

position of the fluorescence maxima blueshifted from 525 to 511 nm. Under these conditions, AB_hRXRB does not have a propensity for LLPS (N), which was confirmed by measurements of turbidity (A_{340}) and DIC (data not shown). Additionally, there were differences in the ANS fluorescence spectra depending on the absence or presence of AB_hRXRB. The major increase in the fluorescence intensity was observed upon the formation of LLPS (P) by AB_hRXRB in the presence of 1.5 M TMAO (Fig. 12B). The position of the fluorescence maxima blueshifted from 525 to 503 nm due to association of the ANS probe with AB_hRXRB. These results can be explained by the structural rearrangement and the altered hydrophobicity of the AB_hRXRB surface in the presence of 1.5 M TMAO. Altogether, the above results indicate that AB_hRXRB LLPS driven by TMAO, but not PEG 8000, induces conformational changes.

Discussion

Nuclear receptors (NRs) are transcription factors (TFs) controlled by small lipophilic ligands [80]. Given the wide variety of genes and processes controlled by NRs, their dysregulation contributes to numerous diseases. As they represent promising therapeutic targets [81], it is crucial to understand the mechanisms underlying their actions and ways in which to selectively modulate their activities. The retinoid X receptor (RXR) is unique among NRs in that it can form dimers with other NRs [15]. In humans, there are three subtypes of RXR (α , β and γ) that differ mostly in their AB regions (Fig. S1). The N-terminal sequence of NRs often lacks a well-defined

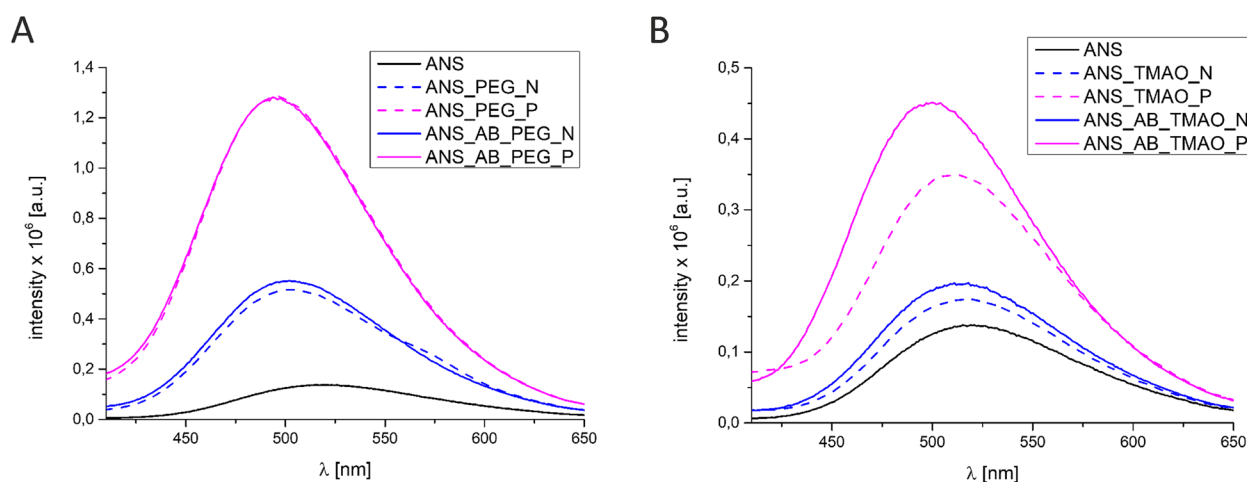


Fig. 12 The effects of PEG 8000 and TMAO on AB_hRXRB. **A** Fluorescence spectra of ANS binding to AB_hRXRB in the presence of 5% (w/v) PEG 8000 (no LLPS (N), blue solid line) and 20% (w/v) PEG 8000 (LLPS (P), magenta solid line). The spectra of free ANS (black solid line) and ANS in the presence of 5% (blue dashed line) and 20% (magenta dashed line) PEG 8000 are also presented. **B** Fluorescence spectra of ANS binding to AB_hRXRB in the presence of 0.5 M TMAO (no LLPS (N), blue solid line) and 1.5 M TMAO (LLPS (P), magenta solid line). The spectra of free ANS (black solid line) and ANS in the presence of 0.5 M (blue dashed line) and 1.5 M (magenta dashed line) TMAO are also presented

three-dimensional structure and possesses properties of IDRs [82, 83]. Each of these sequences has a characteristic disordered pattern that contributes significantly to the specificity of NR action. Upon interaction with coregulators or transcription factors, they may assume a locally stable structure [84]. Subtype β of RXR has the longest AB region, which has not yet been characterized at the molecular level. To describe the biochemical and biophysical properties of the AB region of RXR β (AB_hRXR β), we have elaborated and optimized a protocol for efficient expression and purification. To overcome the problems with the expression of recombinant protein that has a high content of P amino acid residues (25.7%) and enhance its solubility, the SUMO fusion system [29] was chosen. This enabled us to obtain a high level of expression of AB_hRXR β . Additionally, during the purification procedure, the SUMO tag was removed by SUMO protease, resulting in the production of native protein (without additional amino acid residues). High susceptibility to degradation of AB_hRXR β noticed during the purification procedure, an increased electrophoretic mobility and *in silico* results suggested that AB_hRXR β might be an IDR, as was earlier observed for AB_hRXR γ (Table S1) [27]. A comparison of the amino acid composition of both AB regions showed that although they are deficient in order-promoting amino acid residues and abundant in residues characterized as disorder-promoting, the percentage of particular amino acid residues are different (Figure S2). The differences in composition are especially seen for P, S, R, T, Y and N residues. Furthermore, the CD analysis in the far-UV region indicated that AB_hRXR β and AB_hRXR γ have spectra that are characteristic of random coil conformations and confirmed the presence of numerous disordered regions. Despite significant differences in the composition of amino acid residues, the quantitative analysis revealed that $54.5\% \pm 1.7$ of the AB_hRXR β sequence was unordered, which was only slightly lower than the value obtained for AB_hRXR γ (61.8 ± 3.7) (Table S1). None of these values matched the results obtained with the disorder predictors, which indicated disorder of more than 80% in both AB region sequences. The difference between these values may be a consequence of the hidden (not predicted by *in silico* analysis) structure-forming propensity. Both AB regions (AB_hRXR β and AB_hRXR γ) exhibit properties of IDPs/IDRs, but they also possess a substantial but different amount of ordered secondary structure motifs (Table S1). Such a residual structure can be involved in many molecular recognition events and binding to the physiological target of a given protein, which might be important for AB regions function, as they harbor an activation function (AF1) region. The analysis of the hydrodynamic properties of AB_hRXR β and AB_hRXR γ

indicated that, in solution, they exist as monomeric proteins with an extended conformation. However, based on CD and AUC experiments, they were classified into two different groups of IDPs – AB_hRXR β has properties of coil-like IDPs, whereas AB_hRXR γ PMG-like IDPs. These different molecular properties of both sequences might have functional implications e.g., influence LLPS behavior of both proteins.

LLPS propensity depends not only on the presence of particular amino acid residues, but also on the number of residues and their position. The amino acid composition determines the structure, interactions and material properties of proteins, including their propensity to undergo LLPS. For example, LLPS may be driven by cation– π interactions between positively charged residues such as R and K and aromatic side chains such as Y and F, π stacking interactions between aromatic side chains, mainly Y and F [50]. Additionally, G residues increase fluidity, while Q and S residues increase hardening [85]. Although for some amino acid residues it is possible to define the role and the conditions in LLPS phenomenon, for others this can be a challenge. Analysis of the amino acid composition of the AB_hRXR β sequence showed a high content of P amino acid residues. The P amino acid residue has unique physicochemical properties and confers unusual structural properties [86]. It is also one of the most disorder-promoting amino acid residues and is associated with IDRs/IDPs [87]. P-rich motifs (PRMs) are common across the proteomes of many species and appear in various groups of proteins, including cytoskeletal proteins, nucleic acid-binding proteins, transport proteins, and splicing factors [88]. They are common recognition sites for protein–protein interaction modules such as the Src homology 3 (SH3) domain, the WW domain, and the Enabled/VASP homology 1 (EVH1) domain [89]. The AF regions of some TFs (e.g., AP-2 and CTF/NF1) [90, 91] are known to be enriched in P amino acid residues and have been classified on this basis [92]. It has been shown that these regions may serve as adaptor elements that bring together various components of the basal transcriptional machinery and coactivators [92]. Recent studies highlight the importance of PRMs as drivers of LLPS [93–95]. The PRMs of EBNA2 transcription factor has been indicated to play vital role in condensate formation [94]. However, to show the role of P residues, all residues (the content of P residues in the sequence is 27.3%) were mutated. The LLPS of EBNA2, as in the case of AB_hRXR β , is observed in the presence of crowding agent (10% PEG 10 000) and the addition of 1,6-hexanediol significantly reduces droplet formation. On the other hand, higher concentration (300 mM) of NaCl suppresses EBNA2 droplet formation, what was not observed for AB_hRXR β . Although, EBNA2 and AB_hRXR β have high

content of P residues (27.3% and 25.7%, respectively), which probably contributes to conditions of droplet formation, other amino acid residues seem also be important modulators of LLPS. The alignment of the RXR β amino acid sequences of selected vertebrates (*Homo sapiens* (Human), *Canis lupus familiaris* (Dog), *Mus musculus* (Mouse), and *Danio rerio* (Zebrafish) indicates that there is sequence conservation within AB regions of these receptors (Figure S6). Most aligned sequences (except *Danio rerio*) have high content of P residues and contain a polyA stretch in the AB region. These N-terminal sequences of RXR β may exhibit similar intrinsically disordered properties and propensities to LLPS under similar conditions. This indicates that the LLPS propensity is rather conserved among subtypes β of RXR. Moreover, as PRMs can rapidly switch between binding partners, form multivalent complexes [92] and drive LLPS, the high content of P amino acid residues in AB_hRXR β may also be critical for the regulation of gene expression.

The LLPS paradigm provides a new framework to understand the mechanism underlying the physiological function of NRs. NRs that are known to undergo LLPS leading to the formation of MLOs or liquid droplets in vitro include AR [24], ER [26], GR [25], PPAR γ [23] and RXR γ [27, 28]. In different NRs, particular regions or domains with unique biochemical natures present different phase separation potentials. It was previously shown that an intrinsically disordered AB region of hRXR γ (AB_hRXR γ) drives LLPS [27], and the process is ionic strength- and temperature-dependent. The formation of protein LLPS can be mediated by a myriad of different interactions. The sensitivity of AB_hRXR γ LLPS to high ionic strength, kosmotropic salts and 1,6-hexanediol indicates the contribution of hydrophobic interactions in the formation of liquid droplets [28]. Additionally, molecular crowding agents and TMAO enabled phase transition at lower AB_hRXR γ concentrations (Table S1). Here, our experimental data show that AB_hRXR β also has a propensity to induce LLPS. Although the formation of AB_hRXR β droplets is observed in the presence of molecular crowding agents and TMAO, and AB_hRXR β liquid condensates are sensitive to the average molecular mass of crowding agents, as well as sensitive to kosmotropic salts and 1,6-hexanediol, as in the case of AB_hRXR γ , the phase propensity of AB_hRXR β exhibits some unique features. First, in silico analysis indicated that AB_hRXR β may have greater propensity to undergo LLPS than AB_hRXR γ [28]. AB_hRXR β was even classified as a droplet-driver protein. However, our results showed the opposite. AB_hRXR β is not able to form liquid droplets in the

absence of crowding agents in contrast to AB_hRXR γ . Moreover, after addition of crowding agent, the system needed higher concentration of AB_hRXR β and longer time to undergo LLPS. It seems that these AB regions can play a different role in the context of full-length protein. AB_hRXR γ seems to be critical scaffold element for liquid condensates integrity whereas AB_hRXR β seems to be a co-scaffolds region with ability to LLPS only in the presence of another component. Second, AB_hRXR β droplet formation does not depend on the ionic strength or temperature. In the N-terminus of AB_hRXR β , there is a fraction of negatively and positively charged residues. Although, it is possible that there may be electrostatic interactions between positively patches of one molecule of AB_hRXR β and negatively of other, these intermolecular interactions are not sufficient to induce LLPS of AB_hRXR β . Further LLPS analysis confirmed the role of hydrophobic interactions in LLPS of AB_hRXR β and suggested the role of hydrophobic patch located in the middle of the AB_hRXR β sequence (Fig. 9A).

AB_hRXR β also showed a propensity to form liquid condensates in the presence of an osmolyte, TMAO. In contrast to PEGs, TMAO leads to conformational changes in AB_hRXR β what may be relevant to AB_hRXR β function as AB_hRXR β has a potential to interact with other proteins. The main mechanism of TMAO action assumes exclusion of osmolytes from the protein surface [96]. It can also induce the folding of proteins due to the unfavorable interaction of osmolytes with the protein peptide backbone [63]. Molecular dynamic stimulation reported that TMAO may have opposing effects on hydrophobic and charge-charge interactions [97]. Additionally, TMAO may also act as a crowding agent [98] or weaken hydrogen bonds between water molecules and proteins [99]. The CD data indicated changes in AB_hRXR β secondary structure in the presence of TMAO. Additionally, the ANS fluorescence experiments showed that TMAO-induced LLPS triggers changes in the surface hydrophobicity of AB_hRXR β . The observed increase in ANS fluorescence intensity in TMAO-treated AB_hRXR β strongly suggests the exposure of hydrophobic groups and their involvement in LLPS of AB_hRXR β . These observations can be explained by the formation of a polyproline II (PP II) helix that presents an easily accessible hydrophobic surface [100]. Although hydrophobic interactions seem to be important in the formation of liquid droplets by both AB_hRXR β and AB_hRXR γ , the distinct response to the environmental conditions driving LLPS can impact the propensity for droplet formation between a particular subtype of RXR and might be an important element of RXR action.

Conclusions

Transcription is a complex process involving a large number of macromolecules, yet there is still much unknown information. A new aspect of gene expression is the ability of some transcription factors to induce LLPS and MLO formation. RXR is one of 48 human nuclear receptors and is an important ligand-activated transcription factor that has a unique ability to interact with other representatives of the family. Its three subtypes differ mostly in the N-termini (AB regions). Our experimental data show that AB_hRXRβ is a highly disordered region that exhibits a tendency to form locally ordered structures. Additionally, AB_hRXRβ possesses the ability to exhibit LLPS, a phenomenon previously described for AB_hRXRγ. Although both AB regions seem to be similar in terms of their ability to induce phase separation, they clearly differ in the sensitivity to factors driving and regulating LLPS. The distinct LLPS response to environmental factors driven by the unique amino acid compositions of AB_hRXRβ and AB_hRXRγ can have important implications for the transcriptional activity of particular subtypes of RXRs. Further studies will be necessary to disclose how AB regions of RXR, especially in the context of full-length protein, regulate gene expression through phase separation.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12964-023-01113-4>.

Additional file 1: Fig. S1. Alignment of the three human subtypes of RXR amino acid sequences. **Fig. S2.** Percentage of amino acid residues in AB_hRXRβ and AB_hRXRγ. **Fig. S3.** Sedimentation profile of AB_hRXRβ. **Fig. S4.** Das-Pappu phase diagram [43]. **Fig. S5.** CD spectra of AB_hRXRβ in the presence of different factors driving LLPS. **Fig. S6.** Alignment of the RXRβ amino acid sequences of selected vertebrates.

Additional file 2: Table S1. Comparison of results and properties of AB_hRXRβ and AB_hRXRγ.

Acknowledgements

We are grateful to Karolina Partyk (Wrocław University of Science and Technology) for her excellent technical assistance.

Authors' contributions

K.S.: Investigation, Methodology, Project administration, Writing—original draft, Writing—review and editing; A.O.: Funding acquisition, Supervision, Writing—original draft, Writing—review and editing. The authors read and approved the final manuscript.

Funding

The work was supported by a subsidy from the Polish Ministry of Science and Higher Education for the Faculty of Chemistry of Wrocław University of Science and Technology. The mass spectrometry equipment used was sponsored in part by the Centre for Preclinical Research and Technology (CePT), a project cosponsored by the European Regional Development Fund and Innovative Economy, and The National Cohesion Strategy of Poland.

Declarations

Competing interests

The authors declare no conflict of interests.

Received: 10 November 2022 Accepted: 25 March 2023

Published online: 04 May 2023

References

- Mangelsdorf DJ, Borgmeyer U, Heyman RA, Zhou JY, Ong ES, Oro AE, et al. Characterization of three RXR genes that mediate the action of 9-cis retinoic acid. *Genes Dev.* 1992;6:329–44.
- Dawson MI, Xia Z. The retinoid X receptors and their ligands. *Biochim Biophys Acta.* 2012;1821:21–56.
- Röszer T, Menéndez-Gutiérrez MP, Cedenilla M, Ricote M. Retinoid X receptors in macrophage biology. *Trends Endocrinol Metab.* 2013;24:460–8.
- Szanto A, Narkar V, Shen Q, Uray IP, Davies PJA, Nagy L. Retinoid X receptors: X-ploring their (patho)physiological functions. *Cell Death Differ.* 2004;11 Suppl 2:S126–43.
- Weikum ER, Liu X, Ortlund EA. The nuclear receptor superfamily: A structural perspective. *Protein Sci.* 2018;27:1876–92.
- Nocula-Lugowska M, Rymarczyk G, Lisowski M, Ozyhar A. Isoform-specific variation in the intrinsic disorder of the ecdysteroid receptor N-terminal domain. *Proteins.* 2009;76:291–308.
- Pieprzyk J, Zbela A, Jakób M, Ozyhar A, Orłowski M. Homodimerization propensity of the intrinsically disordered N-terminal domain of Ultraspiracle from *Aedes aegypti*. *Biochim Biophys Acta.* 2014;1844:1153–66.
- Dziedzic-Letka A, Rymarczyk G, Kapłon TM, Górecki A, Szamborska-Gbur A, Wojtas M, et al. Intrinsic disorder of *Drosophila melanogaster* hormone receptor 38 N-terminal domain. *Proteins.* 2011;79:376–92.
- Orlov I, Rochel N, Moras D, Klaholz BP. Structure of the full human RXR/VDR nuclear receptor heterodimer complex with its DR3 target DNA. *EMBO J.* 2012;31:291–300.
- Więch A, Rowińska-Żyrek M, Wątyły J, Czarnota A, Hołubowicz R, Szcwczuk Z, et al. The intrinsically disordered C-terminal F domain of the ecdysteroid receptor from *Aedes aegypti* exhibits metal ion-binding ability. *J Steroid Biochem Mol Biol.* 2019;186:42–55.
- Tsafou K, Tiwari PB, Forman-Kay JD, Metallo SJ, Toretzky JA. Targeting Intrinsically Disordered Transcription Factors: Changing the Paradigm. *J Mol Biol.* 2018;430:2321–41.
- Shammas SL. Mechanistic roles of protein disorder within transcription. *Curr Opin Struct Biol.* 2017;42:155–61.
- Gronemeyer H, Gustafsson J-A, Laudet V. Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov.* 2004;3:950–64.
- Pawlak M, Lefebvre P, Staels B. General molecular biology and architecture of nuclear receptors. *Curr Top Med Chem.* 2012;12:486–504.
- Lefebvre P, Benomar Y, Staels B. Retinoid X receptors: common heterodimerization partners with distinct functions. *Trends Endocrinol Metab.* 2010;21:676–83.
- Dollé P, Fraulob V, Kastner P, Chambon P. Developmental expression of murine retinoid X receptor (RXR) genes. *Mech Dev.* 1994;45:91–104.
- Fleischhauer K, McBride OW, DiSanto JP, Ozato K, Yang SY. Cloning and chromosome mapping of human retinoid X receptor beta: selective amino acid sequence conservation of a nuclear hormone receptor in mammals. *Hum Genet.* 1993;90:505–10.
- Tanaka T, De Luca LM. Therapeutic potential of "retinoids" in cancer prevention and treatment. *Cancer Res.* 2009;69:4945–7.
- Kastner P, Mark M, Leid M, Gansmuller A, Chin W, Grondona JM, et al. Abnormal spermatogenesis in RXR beta mutant mice. *Genes Dev.* 1996;10:80–92.
- Mascrez B, Ghyselinck NB, Watanabe M, Annicotte J-S, Chambon P, Auwerx J, et al. Ligand-dependent contribution of RXRβ to cholesterol homeostasis in Sertoli cells. *EMBO Rep.* 2004;5:285–90.
- Zhang R, Fu Z, Fan H, Tian T, Wu M, Xie C, et al. Genetic variant of RXR involved in the vitamin D metabolic pathway was linked to HCV

- infection outcomes among a high-risk Chinese population. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis.* 2021;87:104641.
22. Ishiguro H, Okubo Y, Ohtsuki T, Yamakawa-Kobayashi K, Arinami T. Mutation analysis of the retinoid X receptor beta, nuclear-related receptor 1, and peroxisome proliferator-activated receptor alpha genes in schizophrenia and alcohol dependence: possible haplotype association of nuclear-related receptor 1 gene to alcohol dependence. *Am J Med Genet.* 2002;114:15–23.
 23. Li Z, Luo L, Yu W, Li P, Ou D, Liu J, et al. PPAR γ phase separates with RXR α at PPREs to regulate target gene expression. *Cell Discov.* 2022;8:37.
 24. Ahmed J, Meszaros A, Lazar T, Tompa P. DNA-binding domain as the minimal region driving RNA-dependent liquid-liquid phase separation of androgen receptor. *Protein Sci.* 2021;30:1380–92.
 25. Stortz M, Pecci A, Presman DM, Levi V. Unraveling the molecular interactions involved in phase separation of glucocorticoid receptor. *BMC Biol.* 2020;18:59.
 26. Bojja A, Klein IA, Sabari BR, Dall'Agnesse A, Coffey EL, Zamudio AV, et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell.* 2018;175:1842–1855.e16.
 27. Soltys K, Ozyhar A. Ordered structure-forming properties of the intrinsically disordered AB region of hRXR γ and its ability to promote liquid-liquid phase separation. *J Steroid Biochem Mol Biol.* 2020;198: 105571.
 28. Soltys K, Wycisk K, Ozyhar A. Liquid-liquid phase separation of the intrinsically disordered AB region of hRXR γ is driven by hydrophobic interactions. *Int J Biol Macromol.* 2021;183:936–49.
 29. Butt TR, Edavettal SC, Hall JP, Mattern MR. SUMO fusion technology for difficult-to-express proteins. *Protein Expr Purif.* 2005;43:1–9.
 30. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif.* 2005;41:207–34.
 31. Weeks SD, Drinker M, Loll PJ. Ligation independent cloning vectors for expression of SUMO fusions. *Protein Expr Purif.* 2007;53:40–50.
 32. Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem.* 1989;182:319–26.
 33. Laemmli UK. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature.* 1970;227:680–5.
 34. Weber K, Pringle JR, Osborn M. Measurement of molecular weights by electrophoresis on SDS-acrylamide gel. *Enzym Struct Part C.* Academic Press; 1972. p. 3–27.
 35. Sreerama N, Woody RW. Computation and analysis of protein circular dichroism spectra. *Methods Enzymol.* 2004;383:318–51.
 36. Schuck P, Perugini MA, Gonzales NR, Howlett GJ, Schubert D. Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys J.* 2002;82:1096–111.
 37. Schuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J.* 2000;78:1606–19.
 38. Lebowitz J, Lewis MS, Schuck P. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci.* 2002;11:2067–79.
 39. Fluorescence Polarization. Emission Anisotropy. *Mol Fluoresc.* Wiley; 2001. p. 125–54.
 40. Ferguson RN, Edelhoch H, Saroff HA, Robbins J, Cahnmann HJ. Negative cooperativity in the binding of thyroxine to human serum prealbumin. Preparation of tritium-labeled 8-anilino-1-naphthalenesulfonic acid. *Biochemistry.* 1975;14:282–9.
 41. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy Server. In: Walker JM, editor, et al., *Proteomics Protoc Handb.* Totowa: Humana Press; 2005. p. 571–607.
 42. Vacic V, Uversky VN, Dunker AK, Lonardi S. Composition Profiler: a tool for discovery and visualization of amino acid composition differences. *BMC Bioinformatics.* 2007;8:211.
 43. Holehouse AS, Das RK, Ahad JN, Richardson MOG, Pappu RV. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys J.* 2017;112:16–21.
 44. Mészáros B, Erdős G, Dosztányi Z. IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* 2018;46:W329–37.
 45. Erdős G, Dosztányi Z. Analyzing protein disorder with IUPred2A. *Curr Protoc Bioinforma.* 2020;70:e99.
 46. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. Intrinsic disorder in transcription factors. *Biochemistry.* 2006;45:6873–88.
 47. Cilia E, Pancsa R, Tompa P, Lenaerts T, Vranken WF. From protein sequence to dynamics and disorder with DynaMine. *Nat Commun.* 2013;4:2741.
 48. Cilia E, Pancsa R, Tompa P, Lenaerts T, Vranken WF. The DynaMine web-server: predicting protein dynamics from sequence. *Nucleic Acids Res.* 2014;42:W264–70.
 49. Bolognesi B, Lorenzo Gotor N, Dhar R, Cirillo D, Baldrighi M, Tartaglia GG, et al. A concentration-dependent liquid phase separation can cause toxicity upon increased protein expression. *Cell Rep.* 2016;16:222–31.
 50. Vernon RM, Chong PA, Tsang B, Kim TH, Bah A, Farber P, et al. Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *Elife.* 2018;7:e31486. <https://doi.org/10.7554/eLife.31486>.
 51. Chu X, Sun T, Li Q, Xu Y, Zhang Z, Lai L, et al. Prediction of liquid-liquid phase separating proteins using machine learning. *BMC Bioinformatics.* 2022;23:72.
 52. Hardenberg M, Horvath A, Ambrus V, Fuxreiter M, Vendruscolo M. Widespread occurrence of the droplet state of proteins in the human proteome. *Proc Natl Acad Sci U S A.* 2020;117:33254–62.
 53. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 1997;25:4876–82.
 54. Illiano A, Pinto G, Melchiorre C, Carpentieri A, Faraco V, Amoresano A. Protein glycosylation investigated by mass spectrometry: an overview. *Cells.* 2020;9(9):1986. <https://doi.org/10.3390/cells9091986>.
 55. Tompa P. Intrinsically unstructured proteins. *Trends Biochem Sci Eng-land.* 2002;27:527–33.
 56. Campen A, Williams RM, Brown CJ, Meng J, Uversky VN, Dunker AK. TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. *Protein Pept Lett.* 2008;15:956–63.
 57. Habchi J, Tompa P, Longhi S, Uversky VN. Introducing protein intrinsic disorder. *Chem Rev.* 2014;114:6561–88.
 58. Uversky VN. What does it mean to be natively unfolded? *Eur J Biochem.* 2002;269:2–12.
 59. Manon F, Ebel C. Analytical Ultracentrifugation, a Useful Tool to Probe Intrinsically Disordered Proteins. 2010. <https://doi.org/10.1002/9780470602614.ch15>.
 60. Tcherkasskaya O, Uversky VN. Denatured collapsed states in protein folding: example of apomyoglobin. *Proteins.* 2001;44:244–54.
 61. Kelly SM, Jess TJ, Price NC. How to study proteins by circular dichroism. *Biochim Biophys Acta - Proteins Proteomics.* 2005;1751:119–39.
 62. Uversky VN. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* 2002;11:739–56.
 63. Uversky VN. Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J.* 2009;28:305–25.
 64. Wright PE, Dyson HJ. Linking folding and binding. *Curr Opin Struct Biol.* 2009;19:31–8.
 65. Bah A, Forman-Kay JD. Modulation of intrinsically disordered protein function by post-translational modifications. *J Biol Chem.* 2016;291:6696–705.
 66. Bolen DW, Baskakov IV. The osmophobic effect: natural selection of a thermodynamic force in protein folding. *J Mol Biol.* 2001;310:955–63.
 67. Celinski SA, Scholtz JM. Osmolyte effects on helix formation in peptides and the stability of coiled-coils. *Protein Sci.* 2002;11:2048–51.
 68. Shao Q, Gao YQ. Water plays an important role in osmolyte-induced hairpin structure change: a molecular dynamics simulation study. *J Chem Phys.* 2012;137:145101.
 69. Munishkina LA, Phelan C, Uversky VN, Fink AL. Conformational behavior and aggregation of alpha-synuclein in organic solvents: modeling the effects of membranes. *Biochemistry.* 2003;42:2720–30.
 70. Khan SH, Arnott JA, Kumar R. Naturally occurring osmolyte, trehalose induces functional conformation in an intrinsically disordered activation domain of glucocorticoid receptor. *PLoS ONE.* 2011;6:e19689.
 71. Feng Z, Chen X, Wu X, Zhang M. Formation of biological condensates via phase separation: Characteristics, analytical methods, and physiological implications. *J Biol Chem.* 2019;294:14823–35.
 72. Dignon GL, Best RB, Mittal J. Biomolecular phase separation: from molecular driving forces to macroscopic properties. *Annu Rev Phys Chem.* 2020;71:53–75.

73. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: Organizers of cellular biochemistry. *Nat Rev Mol Cell Biol.* 2017;18:285–98. <https://doi.org/10.1038/nrm.2017.7>.
74. Majumdar A, Dogra P, Maity S, Mukhopadhyay S. Liquid-liquid phase separation is driven by large-scale conformational unwinding and fluctuations of intrinsically disordered protein molecules. *J Phys Chem Lett.* 2019;10:3929–36.
75. André AAM, Spruijt E. Liquid-liquid phase separation in crowded environments. *Int J Mol Sci.* 2020;21(16):5908. <https://doi.org/10.3390/ijms21165908>.
76. Hofmeister F. Zur Lehre von der Wirkung der Salze. *Arch für Exp Pathol und Pharmakologie.* 1888;25:1–30.
77. Sawyer WH, Puckridge J. The dissociation of proteins by chaotropic salts. *J Biol Chem.* 1973;248:8429–33.
78. Shiraki K, Mimura M, Nishinami S, Ura T. Effect of additives on liquid droplets and aggregates of proteins. *Biophys Rev.* 2020;12:587–92.
79. Stryer L. The interaction of a naphthalene dye with apomyoglobin and apohemoglobin: a fluorescent probe of non-polar binding sites. *J Mol Biol.* 1965;13:482–95.
80. Sladek FM. What are nuclear receptor ligands? *Mol Cell Endocrinol.* 2011;334:3–13.
81. Frigo DE, Bondesson M, Williams C. Nuclear receptors: from molecular mechanisms to therapeutics. *Essays Biochem.* 2021;65(6):847–56. <https://doi.org/10.1042/EBC20210020>.
82. Hilser VJ, Thompson EB. Structural dynamics, intrinsic disorder, and allostery in nuclear receptors as transcription factors. *J Biol Chem.* 2011;286:39675–82.
83. Shamilov R, Aneskievich B. Intrinsic disorder in nuclear receptor amino termini: from investigational challenge to therapeutic opportunity. *Nucl Recept Res.* 2019;6:1–16. <https://doi.org/10.32527/2019/101417>.
84. Lavery DN, Mcewan IJ. Structure and function of steroid receptor AF1 transactivation domains: induction of active conformations. *Biochem J.* 2005;391:449–64.
85. Wang J, Choi J-M, Holehouse AS, Lee HO, Zhang X, Jahnel M, et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell.* 2018;174:688–699.e16.
86. Rath A, Davidson AR, Deber CM. The structure of “unstructured” regions in peptides and proteins: role of the polyproline II helix in protein folding and recognition. *Biopolymers.* 2005;80:179–85.
87. Theillet F-X, Kalmar L, Tompa P, Han K-H, Selenko P, Dunker AK, et al. The alphabet of intrinsic disorder: I. Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsic Disord Proteins.* 2013;1:e24360.
88. Williamson MP. The structure and function of proline-rich regions in proteins. *Biochem J.* 1994;297(Pt 2):249–60.
89. Zarrinpar A, Bhattacharyya RP, Lim WA. The structure and function of proline recognition domains. *Sci STKE.* 2003;2003:RE8.
90. Eckert D, Buhl S, Weber S, Jäger R, Schorle H. The AP-2 family of transcription factors. *Genome Biol.* 2005;6:246.
91. Xiao H, Lis JT, Xiao H, Greenblatt J, Friesen JD. The upstream activator CTF/NF1 and RNA polymerase II share a common element involved in transcriptional activation. *Nucleic Acids Res.* 1994;22:1966–73.
92. Frietze S, Farnham PJ. Transcription factor effector domains. *Subcell Biochem.* 2011;52:261–77.
93. Zhang X, Vigers M, McCarty J, Rauch JN, Fredrickson GH, Wilson MZ, et al. The proline-rich domain promotes Tau liquid-liquid phase separation in cells. *J Cell Biol.* 2020;219(11):e202006054. <https://doi.org/10.1083/jcb.202006054>.
94. Peng Q, Wang L, Qin Z, Wang J, Zheng X, Wei L, et al. Phase separation of Epstein-Barr virus EBNA2 and its coactivator EBNA2P controls gene expression. *J Virol.* 2020;94(7):e01771–19. <https://doi.org/10.1042/EBC20210020>.
95. Li P, Banjade S, Cheng H-C, Kim S, Chen B, Guo L, et al. Phase transitions in the assembly of multivalent signalling proteins. *Nature.* 2012;483:336–40.
96. Bruździak P, Adamczak B, Kaczkowska E, Czub J, Stangret J. Are stabilizing osmolytes preferentially excluded from the protein surface? FTIR and MD studies. *Phys Chem Chem Phys England.* 2015;17:23155–64.
97. Su Z, Ravindhran G, Dias CL. Effects of Trimethylamine- N-oxide (TMAO) on Hydrophobic and Charged Interactions. *J Phys Chem B United States.* 2018;122:5557–66.
98. Cho SS, Reddy G, Straub JE, Thirumalai D. Entropic stabilization of proteins by TMAO. *J Phys Chem B.* 2011;115:13401–7.
99. Ma J, Pazos IM, Gai F. Microscopic insights into the protein-stabilizing effect of trimethylamine N-oxide (TMAO). *Proc Natl Acad Sci.* 2014;111:8476–81.
100. Kay BK, Williamson MP, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J Off Publ Fed Am Soc Exp Biol.* 2000;14:231–41.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

