



Yeast recombinant production of intact human membrane proteins with long intrinsically disordered intracellular regions for structural studies

Kassem, Noah; Kassem, Maher M.; Pedersen, Stine F.; Pedersen, Per Amstrup; Kragelund, Birthe B.

Published in:

Biochimica et Biophysica Acta - Biomembranes

DOI:

[10.1016/j.bbamem.2020.183272](https://doi.org/10.1016/j.bbamem.2020.183272)

Publication date:

2020

Document version

Publisher's PDF, also known as Version of record

Document license:

[CC BY](https://creativecommons.org/licenses/by/4.0/)

Citation for published version (APA):

Kassem, N., Kassem, M. M., Pedersen, S. F., Pedersen, P. A., & Kragelund, B. B. (2020). Yeast recombinant production of intact human membrane proteins with long intrinsically disordered intracellular regions for structural studies. *Biochimica et Biophysica Acta - Biomembranes*, 1862(6), [183272]. <https://doi.org/10.1016/j.bbamem.2020.183272>



Yeast recombinant production of intact human membrane proteins with long intrinsically disordered intracellular regions for structural studies



Noah Kassem^a, Maher M. Kassem^b, Stine F. Pedersen^c, Per Amstrup Pedersen^{c,*}, Birthe B. Kragelund^{a,*}

^a Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Ole Maaloes Vej 5, DK-2200 Copenhagen N, Denmark

^b Machine Learning, Department of Computer Science, University of Copenhagen, Universitetsparken 5, DK-2100, Copenhagen Ø, Denmark

^c Cell Biology and Physiology, Department of Biology, University of Copenhagen, Universitetsparken 13, DK-2100 Copenhagen Ø, Denmark

ARTICLE INFO

Keywords:

Growth hormone receptor
GHR
Sodium-proton exchanger 1
NHE1
IDP
Disordered domains
S. cerevisiae
GFP
Nanodiscs

ABSTRACT

Membrane proteins exist in lipid bilayers and mediate solute transport, signal transduction, cell-cell communication and energy conversion. Their activities are fundamental for life, which make them prominent subjects of study, but access to only a limited number of high-resolution structures complicates their mechanistic understanding. The absence of such structures relates mainly to difficulties in expressing and purifying high quality membrane protein samples in large quantities. An additional layer of complexity stems from the presence of intra- and/or extra-cellular domains constituted by unstructured intrinsically disordered regions (IDR), which can be hundreds of residues long. Although IDRs form key interaction hubs that facilitate biological processes, these are regularly removed to enable structural studies. To advance mechanistic insight into intact intrinsically disordered membrane proteins, we have developed a protocol for their purification. Using engineered yeast cells for optimized expression and purification, we have purified to homogeneity two very different human membrane proteins each with > 300 residues long IDRs; the sodium proton exchanger 1 and the growth hormone receptor. Subsequent to their purification we have further explored their incorporation into membrane scaffolding protein nanodiscs, which will enable future structural studies.

1. Introduction

Integral membrane proteins (MP) constitute around 30% of the proteome [1] and they are fundamental for homeostatic maintenance of living cells, which involves a controlled flow of water and substrates across cellular membranes and signal transduction through MP receptors. Knowledge on MP structures has in the past decades been crucial for improving our mechanistic understanding of fundamental cellular processes and for rational drug design [2,3] as MPs constitute > 60% of known drug targets [4]. However, relative to soluble proteins, MP structures are heavily underrepresented in the Protein Data Bank (PDB). Only a small subset (~14%) of known MP structures are of human origin [5,6]. Prokaryotic protein, which are easier to purify due to their shorter extramembrane regions, have consequently been used extensively to model human MPs [7]. This approach is limited by the fact that around 85% of human MPs lack prokaryotic counterparts [8]. The recent technical innovations in cryo-EM [9] and MP crystallography [10] have made structures of human MPs more accessible. Although most MPs that are currently targeted

therapeutically belong to the group of G-protein coupled receptors, they represent only a fraction of the MP continuum. Single pass receptors and multi-spanning MPs, like transporters, pumps and channels, play pivotal roles both physiologically and in pathophysiology, and represent an understudied area in structural biology. This limits development of drugs targeting these MP classes [11]. One reason for their structural absence is likely to be found outside the membrane and is related to their extra-membranous domain(s), of which both folded and intrinsically disordered domains exist. Two human examples where an intact structure is lacking and which harbor significant extra-membranous parts, are the sodium-proton exchanger 1 (hNHE1) [12] and the single-pass cytokine receptor, the growth hormone receptor (hGHR) [13]. hNHE1 contains 12 transmembrane helices but also contains extensive disordered loops and a long, disordered C-terminal tail. hGHR has only 4% of its sequence embedded in the membrane and has a long, disordered intracellular domain. These are issues that complicate recombinant production and purification.

hNHE1 belongs to the SLC9A family and is involved in numerous essential physiological processes [12,14,15]. hNHE1 dysregulation is

* Corresponding authors.

E-mail addresses: PAPedersen@bio.ku.dk (P.A. Pedersen), bbk@bio.ku.dk (B.B. Kragelund).

<https://doi.org/10.1016/j.bbamem.2020.183272>

Received 23 November 2019; Received in revised form 8 March 2020; Accepted 9 March 2020

Available online 10 March 2020

0005-2736/© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

associated with severe diseases including cancer [16], rendering it an interesting therapeutic target. Its primary function is to maintain cytosolic pH homeostasis by 1:1 exchange of extracellular sodium ions for cytosolic protons. hNHE1 activity is tightly regulated through lipid-binding partners, phosphorylation and interactions with multiple proteins through its intrinsically disordered C-terminal tail [12,17,18]. No high-resolution structures exist from any member of the SLC9A family. Models based on bacterial homologs predict a short intracellular N-terminus, 12 transmembrane (TM) helices and a long, partially intrinsically disordered (ID) C-terminal cytoplasmic tail [12], (Fig. 1A).

hGHR is a class 1 cytokine receptor, which regulates key physiological processes such as growth at a cellular and systemic level, metabolism, bone turnover and the immune system [19–21]. It has an extracellular domain (ECD), one TM helix *i.e.* the transmembrane domain (TMD) and a long ID intracellular domain (ICD) (Fig. 1B). Intracellular signaling is initiated by growth hormone binding to the ECD in a 2:1 stoichiometry [22]. Even though the hGHR has been heavily studied for more than three decades, and three-dimensional structures exist of the smaller, globular domains [23,24], no experimentally determined full-length structure exists for this or any other class 1 cytokine-receptor, although a combined experimental and computational model of the prolactin receptor exists [25]. Furthermore, hGHR dysregulation is associated with severe diseases including lung cancer [26], which makes hGHR an important therapeutic drug target.

MPs are notoriously difficult to purify as their native abundance is typically prohibitively low. They are largely understood in terms of well-defined structures and are unlikely to be disordered in a classical sense. However, several MPs contain intrinsic disordered regions (IDRs), often a tail or loop that can be > 300 residues long [12,27–29]. The different physico-chemical properties of the membrane embedded part, the water soluble globular part and the IDR make them difficult to purify and prone to degradation. Moreover, despite the strengths of

cryo-EM and crystallography, these come short in elucidating IDRs within MPs. Thus, MPs with IDRs appear method-orphan in high-resolution structural biology. Instead, low-resolution techniques such as small angle x-ray and neutron scattering (SAXS/SANS) could emerge as a golden standard [30,31]. Studies of MPs within lipid environments have advanced in recent years by developments of the nanodisc technology highly applicable to SAXS/SANS [32,33]. Nanodiscs are derivatives of Apo-A1 containing only its lipid-binding helical-repeat domain [33]. They are amphiphilic membrane scaffold proteins (MSP) that form soluble discoidal particles in the presence of lipids [34,35], with the most utilized MSP being MSP1D1 [32]. Studying MPs in nanodiscs allows a detergent-free environment with no empty detergent micelles complicating scattering data processing.

In light of these unmet needs, we take advantage of a newly developed expression system in the yeast *S. cerevisiae*, which has the capacity to produce large amounts of high-quality MPs [36,37]. This system includes a yeast that features increased plasmid copy numbers from 20 to 200 per cell prior to induction of recombinant protein production, regulated by a galactose inducible promoter. The PAP1500 host strain overexpresses the Gal4 transcription factor concomitantly with recombinant protein production [38]. Target protein genes are codon optimized for *S. cerevisiae* and produced with a C-terminal yeast-enhanced GFP (yeGFP) TEV-yeGFP-His₁₀-tag, suitable for expression and purification optimization. Using this system, a number of human MPs have been purified to homogeneity including aquaporin 10 (AQP10) [39], transient receptor (TRP) channels [40], the chloride channel CLC-1 [41] and the Ether-a-go-go-related gene (hERG) potassium channel [42] with structure determination of human AQP10 (PDB entry 6F7H) [39] and human CLC-1 (PDB entry 6QVC) [41] by x-ray crystallography and cryo-EM, respectively [39,41]. However, common to these structures is that the major parts are fully membrane embedded or is ordered outside of the membrane (AQP10: 84%, CLC-1:

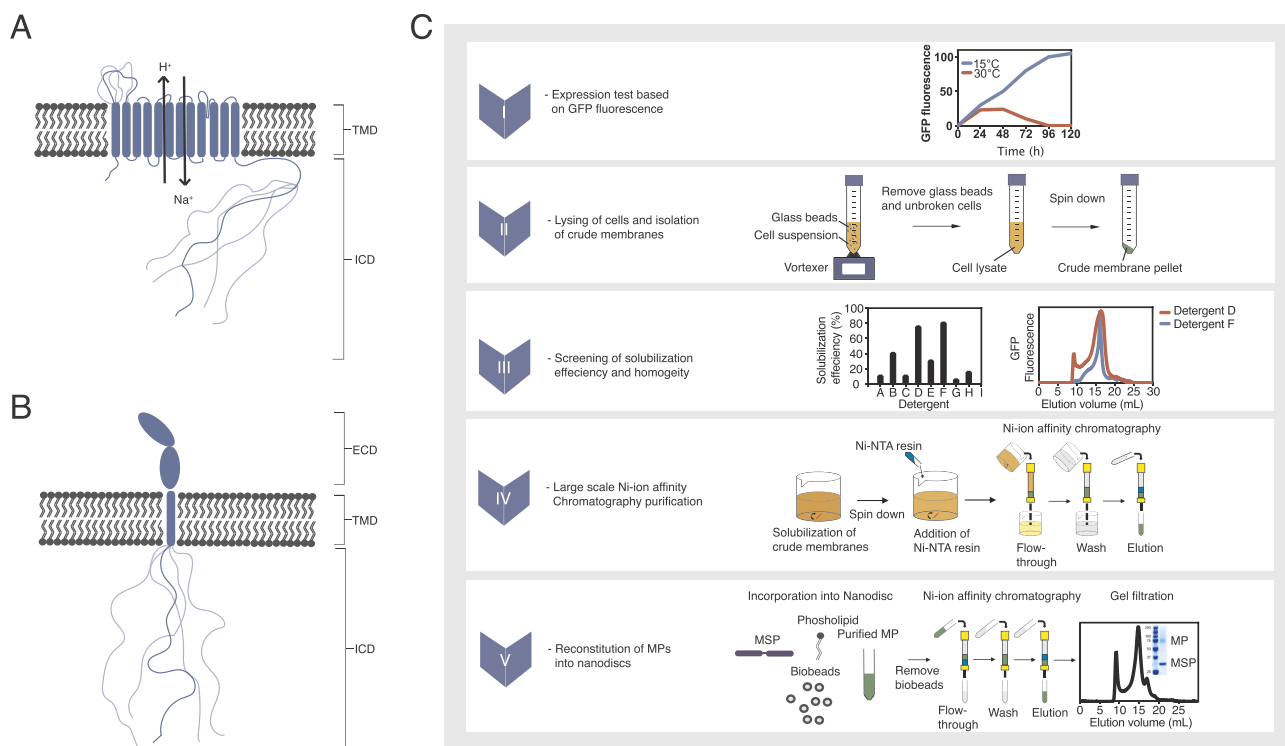


Fig. 1. Cartoons of the human sodium proton exchanger 1 (hNHE1) and the human growth hormone receptor (hGHR). (A) hNHE1 and its 12 transmembrane helices in a lipid bilayer followed by the long C-terminal disordered tail. (B) The full-length hGHR in a lipid bilayer showing its extra-cellular domain, its transmembrane helix and disordered C-terminal tail. (C) Overview of steps from MP expression tests to pure MP incorporated into a nanodisc. The blue arrows from I to V indicate the steps in chronological order. I: Expression test; II: Lysing cells and isolation of crude membranes; III: Detergent screen and FSEC analysis; IV: Large scale purification; V: Incorporation of the MP into nanodiscs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

81%). The aim of this work, therefore, was to provide a suitable protocol for producing intact MPs with long disordered regions for subsequent structural studies. We first survey the abundance and length of disordered regions in human MPs and find that disordered N- and C-terminals appear frequently in the human transmembrane proteome. We then devise a method to express and purify high quality intact human MP including intact long disordered ICDs, using hNHE1 and hGHR as examples (Fig. 1C). In addition to solubilization in detergents, we show that hGHR can be reconstituted into native-like membrane environments with the use of nanodisc technology. Our work lays the framework for future studies of these proteins by SAXS/SANS and cryo-EM, and highlights the need for developing even larger nanodisc systems.

2. Materials and methods

2.1. Disorder prediction of human MPs

Dobson et al. predicts that 26% of the human proteome consists of MPs and annotate the sequences that are experimentally confirmed and predicted sequences with high likelihood of being human MPs. As of September 2019, the human transmembrane protein (HTP) database (<http://htp.enzim.hu>) consists of 5308 sequences [43]. These sequences were applied to the machine learning algorithm Disopred3 (version 3.16) [44]. The algorithm takes a sequence alignment as input. For each sequence we created an alignment using blast (version 2.2.26) [45] with uniref90 as the sequence database [46]. Disopred3 gives a disorder score (between zero and one) per amino acid and labels individual residues “disordered” if the score is equal to or above 0.5 and ordered if below 0.5. To prevent inconsistent segments of ordered and disordered regions, scores were “smoothened” by averaging the scores in a sliding window of 3 in size. After prediction, the number of disordered residues in the N- and C-terminal was counted and grouped.

2.2. Expression system and plasmid constructs

Recombinant hNHE1 and hGHR were produced in *S. cerevisiae* strain PAPI1500 (α *ura3-52 trp1::GAL10-GAL4 lys2-801 leu2Δ1 his3Δ200 pep4::HIS3prb1Δ1.6 R can1 GAL*) [47]. Codon optimized human hNHE1 and hGHR cDNAs for *S. cerevisiae* were purchased from Genscript, USA. Both hNHE1 and hGHR were C-terminally tagged with a Tobacco Etch Virus cleavage site and yEGFP-His₁₀ sequence by PCR amplifying each codon optimized cDNA with the following primers:

hGHR_FW:
5'-TTACTATACTTCTATAGAGACGCAAAGAGAAATAGTAAGATA
ATTATGGATTATGGCAATTATTATTAACCTTAGCTTTAGCTGGTTCT
TCTGATGCT TTTTCTGGTTCAGAAGCTACT-3'

hGHR_RV:
5'-AAATTGACTTTGAAAATACAAATTTTCAGAGTAACCTTCAATG
CAAC-3'

hNHE1_FW:
5'-TTACTATACTTCTATAGAGACGCAAAGAGAAATAGTAAGATAAT
TATGGTTTTGAGATCTGGTATCT-3'

hNHE1_RV:
5'-AAATTGACTTTGAAAATACAAATTTTCTTGACCTTTGGGAAAA
ATGGT-3'

yeGFP_FW
5-GAAAAATTTGTATTTTCAAAGTCAATTTATGAGTAAGGGTGAAGA
ATTATT

yeGFP_RV:
5' GGGCGTGAATCTAAGCGTGACATAACTAATTACATCTAATGGTG
ATGGTGATGGTGATGGTGATGGTGTATATAATTCATCCATACCGT
GAG -3'

For hNHE1 and hGHR the underlined sequences are used for *in vivo* homologous recombination with the expression plasmid and the bold sequence is for recombination with the yeGFP3 [48] PCR fragment to

generate expression plasmids directly in the *S. cerevisiae* expression strain PAPI1500. The bold and underlined sequence is the addition of the hGHR signal peptide. The italicized sequence is the Kozak sequence from the yeast *PMR1* gene. For yeGFP, the bold sequence is for recombination with the anti-codon TEV site. The underlined sequence is used for *in vivo* homologous recombination. The italicized CTA is reverse complementary to the translational stop codon. The sequence in bold font encodes the His₁₀-tag.

2.3. Temperature optimization of hGHR and hNHE1 production

Yeast cells from a glycerol stock were transferred to 5 mL Synthetic Defined (SD) medium (2% (w/v) glucose, 38 mM (NH₄)₂SO₄, 6.4 mM KH₂PO₄, 0.7 mM K₂HPO₄, 2 mM MgSO₄, 1.7 mM NaCl, 81 μM H₃BO₃, 6 μM KI, 27 μM MnSO₄, 43 μM ZnSO₄, 4 μM CuSO₄, 31 μM FeCl₃, 12 μM Na₂MO₄) including leucine (60 mg/L) and lysine (50 mg/L) and grown at 30 °C for 24 h shaking at 120 rpm. An aliquot of 200 μL of the cell culture was transferred to 5 mL SD medium including lysine (50 mg/L) at 30 °C for 24 h. The 5 mL pre-culture was used to inoculate 50 mL of the same medium and grown at 30 °C for 24 h. The 50 mL culture was used to inoculate 2 L of YP medium (1% (w/v) yeast extract 2% (w/v) casein peptone 2% (w/v)) with 3% (v/v) glycerol and 0.5% (w/v) glucose to an OD₄₅₀ = 0.1. The 2 L culture was incubated at room temperature until OD₄₅₀ = 2.5. Half of the culture (1 L) was incubated at 15 °C and the other half at 30 °C. After 30 min, the two cultures were each induced with 110 mL 20% (w/v) galactose dissolved in YP medium containing 3% (w/v) glycerol and no glucose. 50 mL samples were harvested at 0, 24, 48, 72, 96, and 120 h after induction by centrifugation at 4000 g for 20 min. Cells were lysed and crude membranes were isolated at each time point (see below). The GFP fluorescence of 25 μg crude membranes was measured for each timepoint in a spectrofluorimeter (Fluoroskan Ascent, Thermo Scientific). GFP was excited at 485 nm and emission was measured at 520 nm. GFP fluorescence was converted to pmol hNHE1/hGHR from a standard curve generated from purified GFP mixed with yeast membranes as established in [42].

2.4. Crude membrane isolation

1 g yeast cells were mixed with 4 ml glass beads and 1 mL ice cold Lysis buffer (25 mM Imidazole 1 mM EDTA, 1 mM EGTA, 0.5 M KCl, 5% (v/v) glycerol, pH 7.5) containing 1 mM PMSF and 1 μg/ml of leupeptine, pepstatin and chymostatin (LPC) in a 15 mL Falcon tube. The Falcon tube was vortexed at max speed for 1 min and left on ice for 1 min. This was repeated 9 times. The lysed suspension was transferred to a new 15 mL Falcon tube and the beads were washed with 8 mL ice cold Lysis buffer. The suspension was centrifuged at 1000 g for 10 min at 4 °C. The supernatant was ultracentrifuged at 40,000 RPM in a 70Ti rotor for 1.5 h. Crude membranes were resuspended in 0.6 mL Lysis buffer containing freshly added PMSF and LPC and subsequently stored at -80 °C. Protein concentrations in crude membranes were determined by BCA assay [49] following manufacturers' specifications (Sigma, USA) and using chicken ovalbumin as a standard.

2.5. SDS-PAGE and western blotting

SDS-PAGE analysis of protein, in-gel fluorescence and western blotting were performed as described in [42]. A mixture of two rabbit polyclonal anti-GFP-antibodies custom made at Pineda, Germany, was used to detect the hNHE1-GFP and hGHR-GFP fusion proteins in western blots in conjunction with a horse radish conjugated anti-rabbit secondary antibody (Thermo Fisher Scientific). Chemiluminescence was detected using Immobilon Western Chemiluminescence HRP substrate (Millipore *) and the LAS4000 Imager (GE Healthcare, USA).

2.6. Live cell imaging

Localization of heterologously expressed GFP-tagged hNHE1 and hGHR was visualized by GFP fluorescence in whole cells at 1000 × magnification, using a Nikon Eclipse E600 microscope coupled to an Optronics Magnafire model S99802 camera.

2.7. Detergent screens of hNHE1 and hGHR

Crude membranes were incubated in 25 mM Tris-HCl, 10 mM Imidazole, 0.5 M NaCl, 10% (w/v) glycerol, pH 7.5, supplemented with 1 mM PMSF and 1 µg/ml leupeptin, chymostatin and pepstatin, respectively, at a protein:detergent:CHS ratio of 1:3:1. The detergents included in the screen were: Foscholine (FC)-12, n-dodecylphosphocholine; FC-13, n-Tridecylphosphocholine; FC-14, n-Tetradecylphosphocholine; FC-16, n-hexadecylphosphocholine; LDAO, Lauryldimethylamine N-oxide; Cymal-5, 5-cyclohexyl-1-pentyl-β-D-maltoside; DDM, n-dodecyl-β-D-maltopyranoside; DM, n-decyl-β-D-maltopyranoside. The detergents were purchased from Affymetrix, UK or from Glycon Germany and were of anagrade quality. Solubilization was performed at 4 °C for 4 h. Solubilized hGHR/hNHE1 protein was separated from un-solubilized cell debris by ultracentrifugation at 70,000 rpm for 20 min at 4 °C in a Beckman Optima™TLX ultracentrifuge fitted with an S.N. 96U 826 rotor. To detect fluorescence a spectrofluorometer (Fluoroskan Ascent, Thermo Scientific) was used using buffer as background. Samples were excited at 485 nm and emission measured at 520 nm. The efficiency of solubilization was estimated by dividing the GFP-fluorescence of the supernatant by the total fluorescence prior to centrifugation.

2.8. Fluorescence-detection size exclusion chromatography

Solubilized crude membranes were separated on a Superose 6 Increase 10/300 GL column in 20 mM Tris-HCl, 150 mM NaCl and 0.03% (w/v) DDM at a flow rate of 0.5 mL/min. The column was coupled to a Shimadzu Prominence RF-20A fluorescence detector and elution was followed by excitation at 485 nm and measuring emission at 520 nm to visualize the elution profile of GFP and GFP-tagged hNHE1/hGHR.

2.9. Production of hGHR/hNHE1 in the bioreactor

Yeast cells from a glycerol stock were selectively propagated until saturation of 50 mL of SD medium supplemented with 60 mg/ml leucine and 50 mg/ml lysine. Next day, 1 L of SD medium supplemented with 50 mg/ml lysine was inoculated with the 50 ml pre-culture. After overnight growth, the 1 L culture was transferred to 10 L of SD medium supplemented with 20 mg/L adenine, 20 mg/L arginine, 30 mg/L leucine, 60 mg/L lysine, 20 mg/L methionine, 20 mg/L tryptophan, 40 mg/L uracil and with 3% (w/v) glucose and 3% (w/v) glycerol as carbon source and propagated in a 15 L Applikon® bioreactor equipped with an ADI 1030 Bio Controller connected to a PC running the BioExpert® software (Applikon, Holland). The initial part of the fermentation was performed at 20 °C. The bioreactor was fed with glucose to a final concentration of 2% (w/v) when the first amount of glucose had been metabolized. The pH of the growth medium was kept at 6.0 by computer-controlled addition of 1 M NH₄OH. The shift from growth on glucose to growth on glycerol was monitored as a decrease in the rate of NH₄OH consumption. At this point the bioreactor was cooled to 15 °C before induction of recombinant protein production with 1 L of 20% galactose dissolved in the initial growth medium lacking glucose. Yeast cells were harvested after 72 h at 2000 g for 10 min at 4 °C and subsequently stored at -80 °C.

2.10. Purification of full length hGHR

100 g of yeast cells with recombinantly expressed hGHR were

resuspended in 100 mL of ice-cold Lysis buffer containing 25 mM Imidazole 1 mM EDTA, 1 mM EGTA, 10% (v/v) glycerol pH 7.5, 1 mM PMSF and 1 µg/mL of leupeptine, pepstatin and chymostatin (LPC). The resuspended cells were divided into 10 mL fractions in 50 mL Falcon tubes and 15 mL glass beads were added to each fraction. Cells were lysed by vortex mixing for 1 min, with a 1-min rest period on ice. This was repeated 10 times for each tube. The glass beads were washed with 400 mL lysis buffer and the lysate was centrifuged at 4 °C at 2000 g for 10 min. The pellet was discarded, and the supernatant was centrifuged at 40,000 rpm in a 70Ti rotor for 1.5 h. The crude membrane pellet was resuspended in 40 mL lysis buffer. The crude membranes were solubilized by mixing 250 mL 2 × solubilization buffer containing 50 mM Tris pH 7.5, 1000 mM NaCl, 20 mM Imidazole, 1.2% FC-16, 0.2 mM EDTA, 0.2 mM EGTA, 2 mM PMSF, 2 µg/µL LPC with the 40 mL crude membranes and 210 mL ice-cold 18 MΩ water. The crude membranes were solubilized for 9 h at 4 °C. Following centrifugation at 40,000 rpm in a 70Ti rotor for 30 min, the supernatant was transferred to a beaker containing 2 mL Ni-NTA resin (GE Healthcare, USA), pre-equilibrated in 1 × solubilization buffer and left to equilibrate for 8 h at 4 °C. The solution was transferred to a plastic column and the flow-through discarded. The column was washed with 100 CV of 1 × solubilization buffer containing 60 mM Imidazole and 0.0053% (w/v) FC-16 and subsequently eluted using 10 × CV 1 × solubilization buffer containing 300 mM Imidazole and 0.0053% (w/v) FC-16 and in 1 mL fractions. Protein yields were quantified by measuring absorbance at 280 nm, and the seven fractions with the highest absorbance were pooled prior to reconstitution into nanodiscs.

2.11. Purification of full length hNHE1

100 g of yeast cells with recombinantly expressed hNHE1 were resuspended in 100 mL of ice-cold Lysis buffer containing 25 mM Imidazole 1 mM EDTA, 1 mM EGTA, 10% (v/v) glycerol pH 7.5, 1 mM PMSF, 1 mM DTT and 1 µg/mL of leupeptine, pepstatin, and chymostatin (LPC). Cell lysis was performed identically to the hGHR protocol as described above, with the addition of 5 mM DTT. The crude membranes were solubilized by mixing 250 mL 2 × solubilization buffer containing 50 mM Tris pH 7.5, 1000 mM NaCl, 20 mM Imidazole, 0.016% (w/v) FC-16, 0.2 mM EDTA, 0.2 mM EGTA, 2 mM PMSF, 2 µg/µL LPC, 1 mM DTT with the 40 mL crude membranes and 210 mL ice-cold 18 MΩ water. Crude membranes were solubilized for 4 h at 4 °C. Following centrifugation at 40,000 rpm in a 70Ti rotor for 30 min at 4 °C, the supernatant was transferred to a beaker containing 2 mL Ni-NTA resin (GE Healthcare, USA), pre-equilibrated in 1 × solubilization buffer and left to equilibrate for 8 h at 4 °C. The solution was transferred to a plastic column and the flow-through was discarded. The column was washed with 100 CV of 1 × solubilization buffer containing 50 mM Imidazole and 0.0053% (w/v) FC-16 and subsequently eluted using 10 × CV 1 × solubilization buffer containing 300 mM Imidazole and 0.0053% (w/v) FC-16. Protein yields were quantified by measuring absorbance at 280 nm. Selected fractions were concentrated using a Millipore 30,000 MWCO spin filter to 0.4 mL and injected onto a Superose 6 increase 10/300 GL column (GE Healthcare, USA) equilibrated with 20 mM Na₂HPO₄/NaH₂PO₄ (pH 7.4), 150 mM NaCl and 1 mM DTT.

2.12. hGHR incorporation into MSP1D1 nanodiscs

To prepare the lipid stock, POPC (Avanti) dissolved in chloroform was dried as a thin film in a glass tube using a nitrogen stream. Next, the lipid film was solubilized in 20 mM Tris, pH 7.5, 100 mM NaCl, 100 mM cholate (cholate buffer) to a final concentration of 50 mM POPC. hGHR with POPC and MSP1D1 in a final ratio of 1:70:20 was prepared as follows: hGHR eluted from the Ni-column was concentrated using a Millipore 30,000 MWCO spin filter to 2 mL and diluted with 20 mM Tris, pH 7.5, 100 mM NaCl, 1 × CMC FC-16 to 15 mL, to reach a

final Imidazole concentration of 40 mM. The sample was then concentrated to 2 mL giving a hGHR concentration of 20 μ M, and was mixed with POPC dissolved in cholate buffer and left to incubate on ice for 15 min. MSP1D1 was added to the mixture, which was then left to incubate for 5 min on ice prior to adding freshly prepared bio-beads SM2 (Biorad) (700 μ L beads per 1000 μ L sample). Bio-Beads bind to detergent effectively removing them from the system thereby initiating the reconstitution of hGHR into MSP1D1. The sample was incubated for 12 h at 4 °C on a tilting table. To remove the Bio-Beads, the sample was transferred into Eppendorf tubes, punching a small hole in the bottom of each of the Eppendorf tubes with a needle, and centrifuged at 500 g for 2 min into 15 mL Falcon tubes, thereby leaving the beads dry in the Eppendorf tubes and the sample in the 15 mL Falcon tubes. The sample was diluted 4 times in 20 mM Tris, pH 7.5, 100 mM NaCl, and 2 mL equilibrated Ni-resin was added to the sample. The sample was allowed to bind to the Ni-resin for 4 h. The sample and Ni-resin were transferred to a plastic column, and the flow-through containing MSP1D1 discs without MP was collected. The Ni⁺-column was washed with 5 CV 20 mM Tris, pH 7.5, 100 mM NaCl, 10 mM Imidazole and then eluted with 10 mL 20 mM Tris, pH 7.5, 100 mM NaCl, 400 mM Imidazole. The entire 10 mL were concentrated using a Millipore 30,000 MWCO spin filter to 0.4 mL and injected onto a Superose 6 increase 10/300 GL column (GE) equilibrated with 20 mM Tris, pH 7.5, 100 mM NaCl. The fractions corresponding to the reconstituted hGHR were assessed by SDS-PAGE.

2.13. MSP1D1 and MSP1E3D1 protein expression and purification

pET28(+) encoding the His-tagged MSP1D1 gene was kindly provided by Professor Lise Arleth (University of Copenhagen) and the plasmid for production of MSP1E3D1 was from Addgene. The proteins were purified using the same protocol. First, the plasmid was transformed into *E. coli* BL21 (DE3) by adding 1 μ L 100 ng/ μ L DNA to 100 μ L of competent cells and plated on a LB-agar plate containing 100 mg/L kanamycin and grown at 37 °C overnight. The following day, a single colony was transferred to 10 mL of LB medium containing 100 mg/L kanamycin and grown for 12 h at 37 °C. The next day, the 10 mL cultures were transferred to 1 L medium in a 5 L flask and grown at 37 °C until the optical density at 600 nm reached 0.6. The culture was subsequently induced with 1 mM IPTG for 3 h prior to harvesting the culture at 4000 g. The cells were resuspended in 40 mL buffer A (50 mM Tris-HCl, pH 8, 300 mM NaCl, 20 mM Imidazole, 6 M GuHCl) and lysed by sonication on ice for 60 s at 90% amplitude using a UP400S Ultrasonic Processor. The suspension was centrifuged at 20,000 g for 15 min and the pellet was discarded. 10 mL of Ni-NTA resin (Qiagen, Germany) equilibrated in buffer A and was added to the supernatant and incubated for 1 h at room temperature. The supernatant was loaded onto a plastic column and the flow-through was discarded. The column was washed in 3 CVs of buffer A, 3 CVs of wash buffer (50 mM Tris-HCl, pH 8, 300 mM NaCl, 40 mM Imidazole) including 10 mM sodium cholate, 3 CVs of wash buffer, and protein was eluted in buffer B (50 mM Tris-HCl, pH 8, 300 mM NaCl, 400 mM Imidazole). Tobacco etch virus (TEV) protease was added to the protein sample at 1:100 mass ratio, and the mix was dialyzed 100-fold against TEV buffer (50 mM Tris-HCl, pH 8, 100 mM NaCl, 1 mM EDTA, 1 mM DTT) overnight at 4 °C, producing N-terminally cleaved MSP1D1/MSP1E3D1 without the His-tag. MSP1D1/MSP1E3D1 was separated from the His-tagged TEV protease and residual His-tagged MSP1D1/MSP1E3D1 by adding Ni-NTA resin to the sample. The MSP1D1/MSP1E3D1 sample was then dialyzed 100-fold in 20 mM Tris-HCl, pH 7.4, 100 mM NaCl overnight at 4 °C.

2.14. Microscale thermophoresis (MST)

hGH_{G120R} was purified as in [50] and human prolactin (hPRL) purified as in [51] and were both labeled with NT-647-NHS [52] using

the Monolith NT™ Protein Labeling Kit RED-NHS (NanoTemper Technologies) for 1 h at room temperature with NT-647-NHS at a molar ratio of 1:3 in labelling buffer following the protocol. These conditions favor modification of the N-terminal amino group. Free dye was separated from reacted dye using the provided desalting column. The ratio between fluorophore and protein was 0.8 for hGH_{G120R} and 0.6 for hPRL. The raw fluorescence change was used to determine the binding affinity. A two-fold dilution series of monomeric hGHR from 750 nM to 23 pM was prepared in 20 mM Na₂HPO₄/NaH₂PO₄ (pH 7.4), 100 mM NaCl with either 20 nM hGH_{G120R} or 20 nM hPRL in each sample. hGH_{G120R} was measured in triplicates whereas hPRL was measured once. Samples were loaded into Monolith NT.115 Premium Capillaries (NanoTemper Technologies), and thermophoresis and raw fluorescence signals measured at 25 °C with a light-emitting diode (LED) power of 80% and an infrared (IR) laser power of 100%. For hGH_{G120R}, the dissociated constant K_D was obtained by fitting the data to $Y = Y_0 + (Y_F - Y_0) / (2 * [P]_{total}) * (K_d + [P]_{total} + X - \sqrt{((K_d + [P]_{total} + X)^2 - 4 * [P]_{total} * X)})$, where Y is the measured fluorescence/MST, X is the ligand, $[P]_{total}$ is the total concentration of the protein, Y_F is the estimated end point of the titration and Y_0 is the start point.

2.15. Circular dichroism

Far-UV CD spectra were recorded on 2 μ M hNHE1 in 0.0053% FC-16 and 2 μ M hGHR in 0.0053% FC-16. The spectra were recorded on a Jasco J-810 Spectropolarimeter in a 1 mm Quartz Suprail cuvette (Hellma) at 20 °C. 5 scans were accumulated from 260 nm to 190 nm for both NHE1 and GHR. The scan mode was continuous with 20 nm/min and a data pitch of 0.5 nm. Buffer was recorded identically and subtracted. The spectra were processed and smoothed (mean-movement methods, convolution 10) and then converted into mean residue ellipticity values using the formula $[\theta] = mdeg / (10 * c * n * l)$. To calculate the expected secondary structure content in the absence of a full-length structures, the Nygaard model [53] was used as the model for NHE1 and the crystal structure of GFP (1GFL) [54] included. To determine the expected secondary structure content for hGHR, the crystal structure of hGHR [23] ECD, the NMR structure of the TMD [55], characterization of the disordered ICD by CD and NMR spectroscopy [29] and GFP (1GFL) [54] was used. To determine the secondary structure for the purified hNHE1 and hGHR the CD data was analyzed by BeStSel [56] using standard settings.

3. Results

3.1. Prevalence of disordered tails in human MPs

We applied bioinformatics to quantify the occurrence of N- and C-terminal IDRs in human MPs. The database we used contains 5308 annotated α -helical MP sequences with high likelihood (> 98%) of being of human origin [43]. The sequences were analyzed using Disopred3 [44] and the histograms in Fig. 2 categorize the MP proteome with respect to length and position of the IDR. We find that ~22% of the MPs do not have a disordered N-terminal longer than five residues (Fig. 2A) and ~51% do not have a disordered C-terminal longer than five residues (Fig. 2B). However, when we combine these, only ~12% of the human MPs are completely without terminal IDRs (< 5 residues) (Fig. 2C). In fact, MPs with a disordered N-terminal tail longer than 25 residues account for ~35% of the MPs with an average length of 54 residues. For C-terminal disorder, ~24% of MPs have IDRs with an average length of 75 residues, allocating as much as 2813 MPs to this category. Even more remarkable, roughly 10% of all human MPs have long disordered regions (> 100 residues) in either the N- or the C-terminal counting over 320 human MPs. With much of the associated biology being linked to these disordered regions, it is striking how disorder is underrepresented in structural studies of MPs. Thus, there is a need to develop protocols for their isolation and for structural studies.

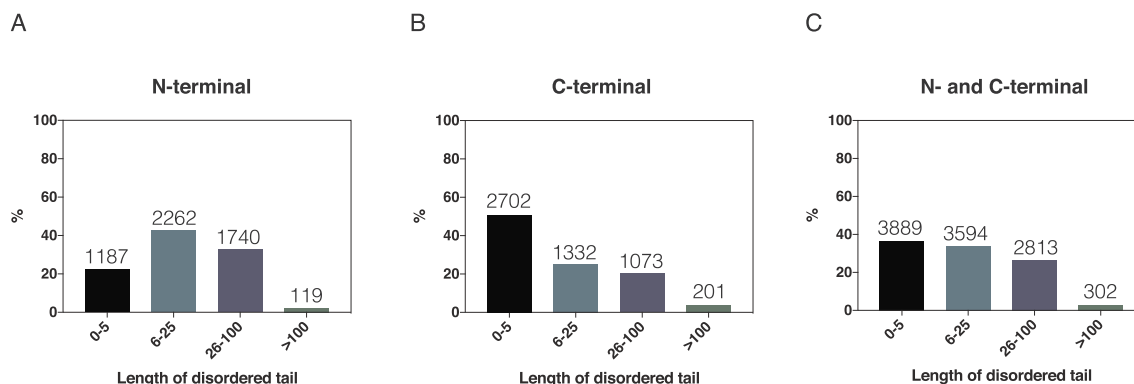


Fig. 2. Disorder predictions of the human transmembrane proteome. The percentage of sequences in a library of 5308 MPs containing between, 1–5, 6–25, 26–100 and > 100 disordered residues are shown for (A) MPs with N-terminal disorder, (B) C-terminal disorder and (C) Either N- or C-terminal disorder. The total number of MPs for each category is listed above the columns. Disorder was predicted using Disopred3 [44].

3.2. hGHR and hNHE1 localize differently in *S. cerevisiae*

The most prominent problem reflected in the lack of MP structures with extensive IDRs, is related to sample preparation. To monitor the expression of hNHE1 and hGHR in PAP1500 cells, we used the fluorescence from the GFP tag. In humans, hNHE1 is expressed in essentially all cells, localizes to the plasma membrane [12] and is both N- and O-glycosylated [57]. The hGHR is expressed in almost all cells [58]. It is synthesized and folded as a precursor protein in the endoplasmic reticulum. hGHR N-glycosylation at five different positions is initiated in the ER and finalized in the Golgi apparatus [59], resulting in a mature receptor that is transported to the plasma membrane [60]. In *S. cerevisiae*, hNHE1 was primarily present in the plasma membrane (Fig. 3A). However, in comparison to the expression of free GFP, which distributed uniformly across the cell cytosol (Fig. 3B), hGHR (Fig. 3C) accumulated in internal membranes, which could be the result of overexpression. Furthermore, previous studies in the same expression system have demonstrated that other MPs purified from intracellular membranes are functional [36,40,42]. Therefore, we continued optimizing expression levels for both MPs.

3.3. Accumulation is favored at a low temperature

We first used the fluorescence from the GFP-tag to measure the accumulation of hNHE1 and hGHR in yeast crude membranes at 15 °C and 30 °C. hNHE1 accumulated to a much higher membrane density at 15 °C compared to 30 °C, reaching a density of 0.7% (w/w) in the crude membranes 72 h after induction at 15 °C (Fig. 4A). The maximum amount of accumulated hNHE1 was 3 times higher at 15 °C compared to 30 °C. The expression of hGHR at 15 °C peaked at a hGHR density of 0.3% (w/w) in crude membranes after approximately 72 h (Fig. 4B). The accumulation leveled off between 72 and 120 h. hGHR also accumulated to almost the same degree after 24 h at 30 °C as at 15 °C after 72 h. The maximal amount of accumulated hGHR was only 1.2 times higher at 15 °C compared to 30 °C.

To investigate the temperature dependent accumulation demonstrated in Fig. 4A and B, we took advantage of the GFP-tag as it has previously been shown in *E. coli* that a C-terminal GFP tag only folds correctly *in vivo* if the membrane embedded MP fusion partner folds correctly [61]. This is based on the observations that GFP fluorescence is preserved after SDS-PAGE analysis and correctly folded GFP only contributes with 10–15 kDa to the molecular weight of its denatured MP fusion partner [61]. Fully denatured GFP increases its apparent SDS-PAGE derived molecular weight to approximately 28 kDa. Thus, only correctly folded GFP is detected by in-gel fluorescence while both correctly folded and incorrectly folded GFP can be visualized by western blotting. We subsequently analyzed the hNHE1 and hGHR content

in crude membranes by in-gel fluorescence (Fig. 4C and D) and western blotting using an anti-GFP-antibody (Fig. 4E and F). From the in-gel fluorescence, it was evident that hNHE1 migrated to the apparent mass of approximately 110 kDa (Fig. 4C), and that this band correlated with the band at 110 kDa on the western blot (Fig. 4E). This indicated correctly folded hNHE1 at both temperatures, but with higher accumulation at 15 °C. A band at approximately 70 kDa was visible in both the in-gel-fluorescence and the western blot, which indicated the presence of a degradation product, but no larger bands were visible indicating absence of aggregated hNHE1 (Fig. 4C and E). From the in-gel fluorescence and western blotting of hGHR-GFP it was evident from hGHR expressed at 15 °C that a band with an apparent mass of approximately 125 kDa could be detected, but also a band between the stacking and running gel (Fig. 4D and F). This may indicate the presence of aggregated hGHR, in accordance with the overexpression observed. The ratio between correctly folded hGHR and presumed aggregated hGHR was higher when expressed at 30 °C compared to 15 °C. Thus, both proteins could be expressed in a folded state in yeast and were present in the crude membrane fraction when extracted.

3.4. Detergent screens and FSEC screening for optimized purification

To purify hNHE1 and hGHR, identification of a detergent suitable for solubilization and purification is essential, and considerable effort should be allocated to screen for conditions that favor both solubility and homogeneity [62]. To identify conditions that efficiently solubilized hNHE1 and hGHR, we set up a screen with eight different detergents chosen to include different hydrophobicities, critical micelle concentrations, and non-ionic and zwitterionic moieties (Table 1). The non-ionic detergents and lauryl-dimethylamine N-oxide (LDAO) were also included due to their successful contribution to MP crystallography [63,64]. It is still debated whether Foscholines (FCs) can act as chaotropic agents, but recent studies have shown that MPs purified in FCs can maintain activity [36,42]. The screen was performed at a protein to detergent ratio of 1:3 in the presence and absence of cholesteryl-hemisuccinate (CHS) previously shown to increase sample homogeneity [42]. The solubilization efficiency, which we define as the GFP fluorescence post-solubilization divided by GFP fluorescence pre-solubilization, was used to quantify solubility of the MPs in detergent. hNHE1 and hGHR showed better solubilization in the zwitterionic FCs (Fig. 5). The most effective detergent for both hNHE1 and hGHR was FC-16, with solubilization efficiencies of approximately 90% and 60%, respectively (Fig. 5).

Solubilization efficiency does not necessarily correlate with the homogeneity of the sample. Therefore, both hNHE1 and hGHR were analyzed by fluorescence detected size exclusion chromatography [65] (FSEC) in all tested detergents with and without CHS (Fig. 5). None of

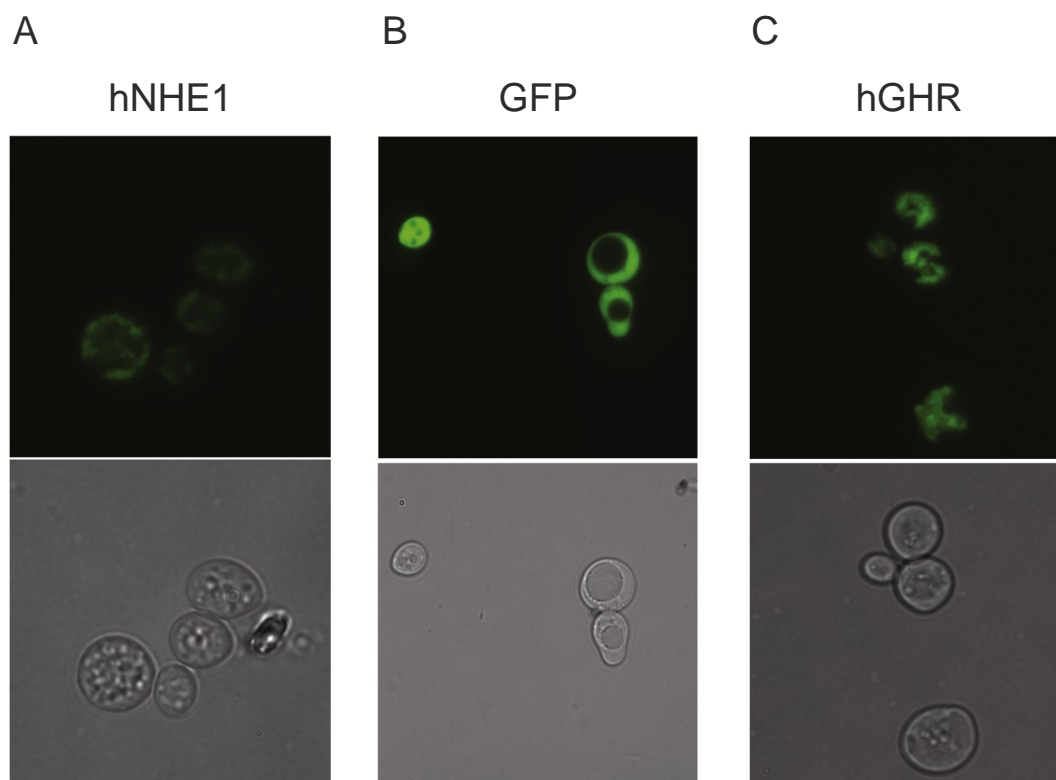


Fig. 3. Live cell imaging of yeast cells producing hNHE1, hGHR or GFP. (A) *S. cerevisiae* cells producing hNHE1-TEV-GFP-His₁₀, (B) free GFP or (C) hGHR-TEV-GFP-His₁₀. *S. cerevisiae* cells were grown at room temperature in expression medium and induced for 72 h at 15 °C as described in [Materials and methods](#). GFP and DIC images are shown pairwise.

the non-ionic detergents (DM, DDM, Cymal-5) were able to solubilize NHE1, but only free GFP eluting at approximately 18 ml was visible ([Fig. 6A-C](#)). Upon solubilization of hNHE1 in zwitter-ionic detergents (LDAO, FC-12, FC-13, FC-14, FC-16), a peak at an elution volume at approximately 15 ml was observed corresponding to recombinantly expressed hNHE1 ([Fig. 6D-H](#)). Compared to all tested detergents, FC-16 was most effective both in regard to solubilization and homogeneity of hNHE1. The hGHR was also mainly soluble in the zwitter-ionic detergents and the homogeneity of the recombinantly expressed hGHR was, similarly to hNHE1, optimal in FC-16 ([Fig. 6I-P](#)). The effect of adding CHS was for both hNHE1 and hGHR a decrease in the amount of solubilized MP. However, as seen by [Fig. 7A-D](#), and [Fig. 7I-L](#), the addition of CHS resulted in larger peaks at an elution volume of approximately 15 ml, but it was not shown to have significant effects for the FCs. Therefore, it was chosen not to use CHS for purification and continue with FC-16 for solubilization.

3.5. Circular dichroism shows folding of the recombinant MPs

The C-terminal His₁₀-tag was exploited for Ni-affinity chromatography immediately after solubilization of both hGHR and hNHE1 in FC-16. To determine the purity of hGHR and hNHE1, we separated the elution by SDS-PAGE and visualized the hGHR and hNHE1 content by in-gel fluorescence and subsequently by Coomassie staining ([Fig. 7A](#)). This showed that the purity after Ni-affinity chromatography was high as no or only few non-fluorescent bands were visible. To further purify hGHR and hNHE1, a Superose 6 increase 10/300 SEC column (GE Healthcare) was used ([Fig. 7B](#)). The SEC analysis revealed that the Ni-affinity chromatography purified samples of hNHE1 and hGHR were homogenous in FC-16 ([Fig. 7B](#)). After SEC, the protein purity was high and even fewer non-fluorescent bands were visible after Coomassie staining. Furthermore, we calculated the purity of the individual fractions during the purification by calculating the amount of GFP tagged

protein from the GFP standard curve and comparing with the amount of purified protein as determined by the BCA assay ([Table 2](#)). [Table 2](#) shows that we purified 0.59 mg hNHE1 per liter of cell culture and 0.30 mg hGHR per liter of cell culture with purities of 84% and 88%, respectively. As some fluorescent bands are still present suggesting a minor population of degradation products, the purity may be slightly overestimated. Still, since IDPs bind less Coomassie than BSA [[66,67](#)] the content of full-length GHR is underestimated.

However, as FC-16 is not the most commonly used detergent for extraction of MPs from crude membranes, the possibility of unfolding of hNHE1 and hGHR exists. We therefore applied far-UV circular dichroism (CD) spectroscopy to verify that the purified proteins attain the expected content of secondary structures ([Fig. 7C](#)). The CD spectra revealed that NHE1 primarily contains an α -helical secondary structure with minima at 208 nm and 222 nm in line with the suggested model [[68](#)] ([Fig. 7C](#), left). From the solved partial structures of hGHR [[23,24](#)], we expect a combination of β -sheet secondary structures and disorder, which is confirmed by the CD spectrum ([Fig. 7C](#), right). Furthermore, preservation of the fluorescence from the C-terminal GFP tag strongly supports that both hNHE1 and hGHR are folded. To further substantiate this, we used BeStSel [[56](#)] to extract the secondary structure content from the experimental CD data and compared this to expected secondary content from available models and crystal structures (see [Materials and methods](#)), [Table 3](#) and [Fig. 7C](#). In the comparison, it is evident that the α -helix content of hNHE1 obtained from the model (27%) fits well with the extracted values from the CD data (23%). The model also included 11% β -sheet structures, which fits well with 9% from the GFP tag. For GHR, and despite the contribution from aromatic exciton coupling in the ECD of the hGHR to the far-UV CD spectrum, the α -helix content in hGHR from the model (8%) again fits well with the extracted values from the CD data (11%). The β -sheet content from the CD data (22%) had a slightly larger discrepancy (29%), which could be attributed to the aromatic exciton coupling from the ECD.

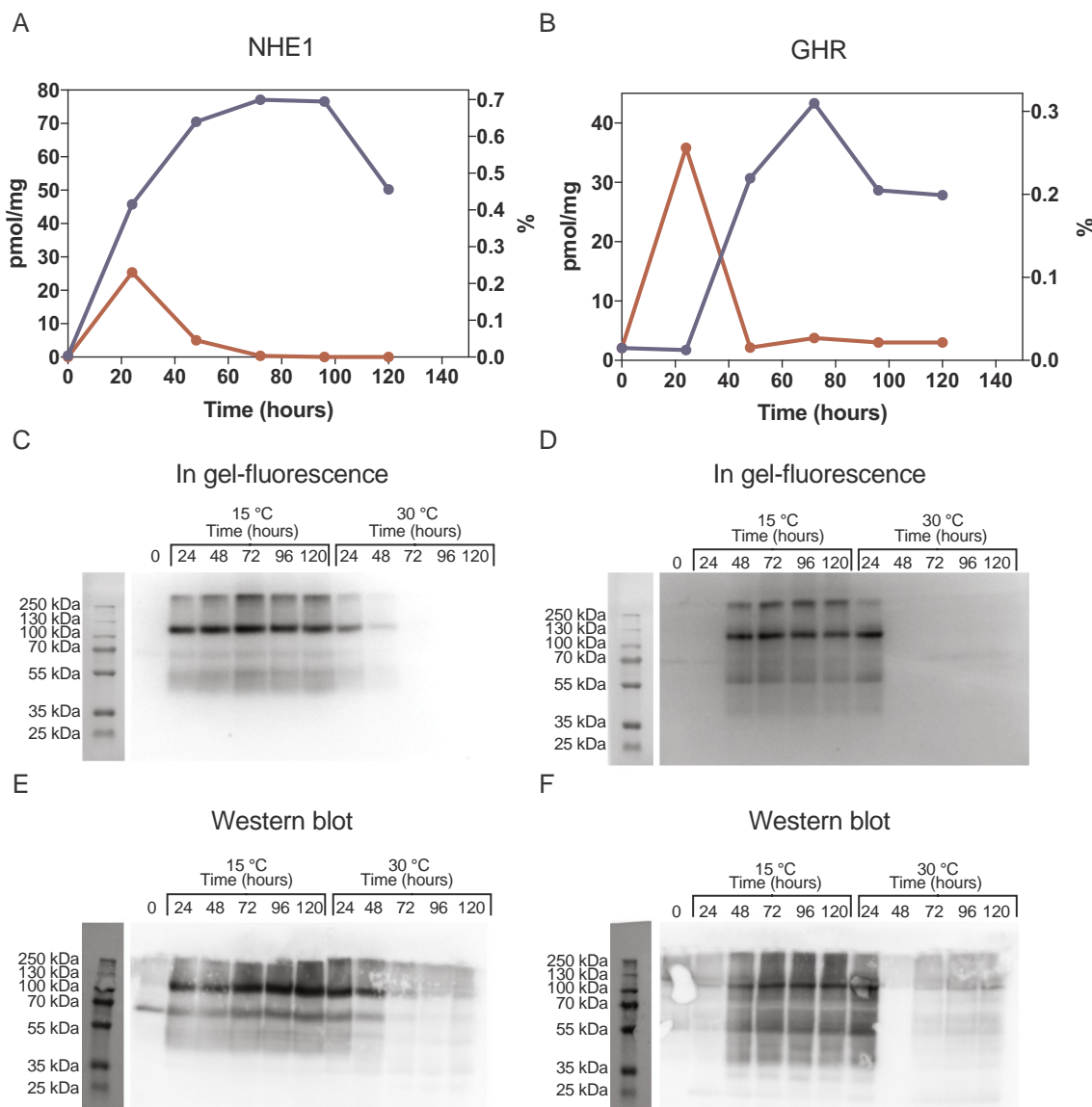


Fig. 4. Accumulation of hNHE1-TEV-GFP-His₁₀ and hGHR-TEV-GFP-His₁₀ in *S. cerevisiae* crude membranes during production at 15 °C or 30 °C. (A, B) *S. cerevisiae* cells were grown at room temperature until OD₄₅₀ = 2.5 and the cultures were divided in two. One half was transferred to 15 °C (blue) and the other was transferred to 30 °C (red). After thermo-equilibration hNHE1 or hGHR production was initiated by the addition of 2% galactose. Crude membranes were isolated from cells at 0, 24, 48, 72, 96 and 120 h after induction and fluorescence was measured in 25 µg crude membranes. This was converted to pmol hNHE1-GFP fusion (A) or pmol hGHR-GFP fusion (B) per mg protein in the crude membranes as described in [Materials and methods](#). This was subsequently converted to the percentage (w/w) of hNHE1 or hGHR in the crude membranes. (C, D) From each timepoint at both 15 °C and 30 °C 25 µg crude membranes were analyzed by SDS-PAGE for both (C) hNHE1 and (D) hGHR and were visualized by in-gel fluorescence. (E, F) After fluorescence detection, the MPs were blotted onto a PVDF membrane and both (E) hNHE1 and (F) hGHR were visualized with an anti-GFP antibody. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Properties of used detergents.^a

| Detergent | Ionic state | Molecular weight (g/mol) | Aggregation number | CMC (H ₂ O) |
|-----------|---------------|--------------------------|--------------------|-----------------------------|
| DM | Non-ionic | 482.6 | ~69 | 1.8 mM (0.087% (w/v)) |
| DDM | Non-ionic | 510.6 | ~149 | 0.17 mM (0.0087% (w/v)) |
| CYMAL5 | Non-ionic | 494.5 | ~47 | 5 mM (0.12% (w/v)) |
| LDAO | Zwitter-ionic | 229.4 | ~76 | 1–2 mM (0.023–0.046% (w/v)) |
| FC12 | Zwitter-ionic | 351.5 | ~54 | ~1.5 mM (0.047% (w/v)) |
| FC13 | Zwitter-ionic | 365.5 | ~87 | ~0.75 mM (0.027% (w/v)) |
| FC14 | Zwitter-ionic | 379.5 | ~108 | ~0.12 mM (0.0046% (w/v)) |
| FC16 | Zwitter-ionic | 407.5 | ~178 | ~0.013 mM (0.00053% w/v) |

^a Aggregation numbers and CMCs taken from <http://www.anatrace.com>

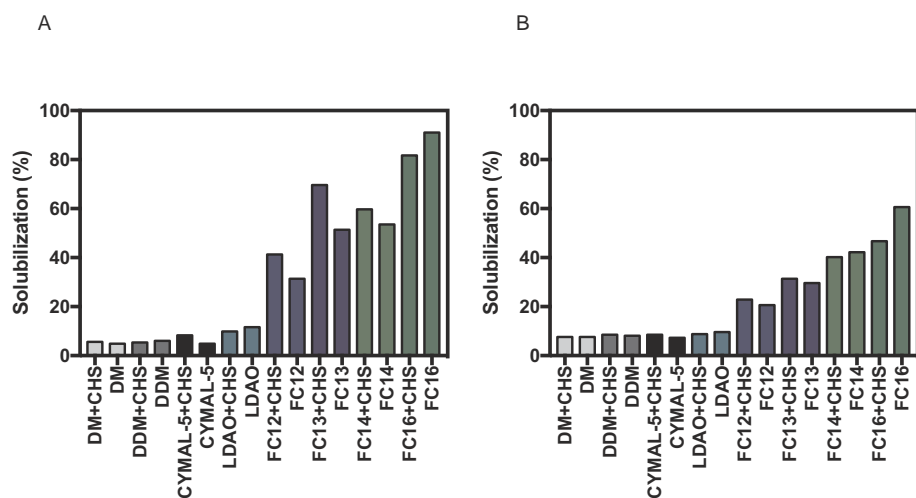


Fig. 5. Detergent screen for (A) hNHE1 and (B) hGHR solubilization. hNHE1 and hGHR were solubilized at a 3:1 detergent to protein ratio (w/v) as described in **Materials and methods**. Solubilization experiments were performed in the absence and presence of cholesterol-hemi-succinate at a 3:1 protein to detergent ratio. Abbreviations; DM, n-Decyl- β -D-maltopyranoside; DDM, n-Dodecyl- β -D-maltopyranoside; Cymal-5, 5-Cyclohexyl-1-pentyl- β -D-maltoside; LDAO, lauryldimethylamine N-oxide; FC, Fos-Choline; CHS, cholesterol-hemi-succinate. Solubilization efficiencies were calculated as the GFP fluorescence of the solubilized samples after ultracentrifugation, divided by the GFP fluorescence before ultracentrifugation.

In conclusion, we present a protocol that is able to produce intact MPs with large disordered intracellular domains to homogeneity and in adequate amounts for structural analyses. To enable such studies, we proceeded to investigate if reconstitution into nanodiscs from FC-16 would be feasible.

3.6. Reconstitution of hGHR and hNHE1 into nanodiscs

Incorporation of MPs into nanodiscs has emerged as a reliable technique for studying MPs in a more native-like lipid bilayer environment void of detergents and in solution for both cryo-EM and

solution scattering techniques such as SAXS and SANS [69–72]. What these techniques have in common is that they allow studies of MPs in solution and reconstituted from the conventional golden standard detergent, DDM. In this study, DDM was not an option due to the low solubility of both hGHR and hNHE1 and instead we used FC-16. The most structurally characterized and utilized MSP for nanodisc reconstitution is the MSP1D1 [73] and the most abundant lipids in the plasma membrane are the phospholipids. POPC is a synthetic phospholipid that is typically used for nanodisc formation [71,73]. We therefore chose this combination for reconstitution. The reason for choosing nanodiscs was two-fold. First, nanodiscs are readily available

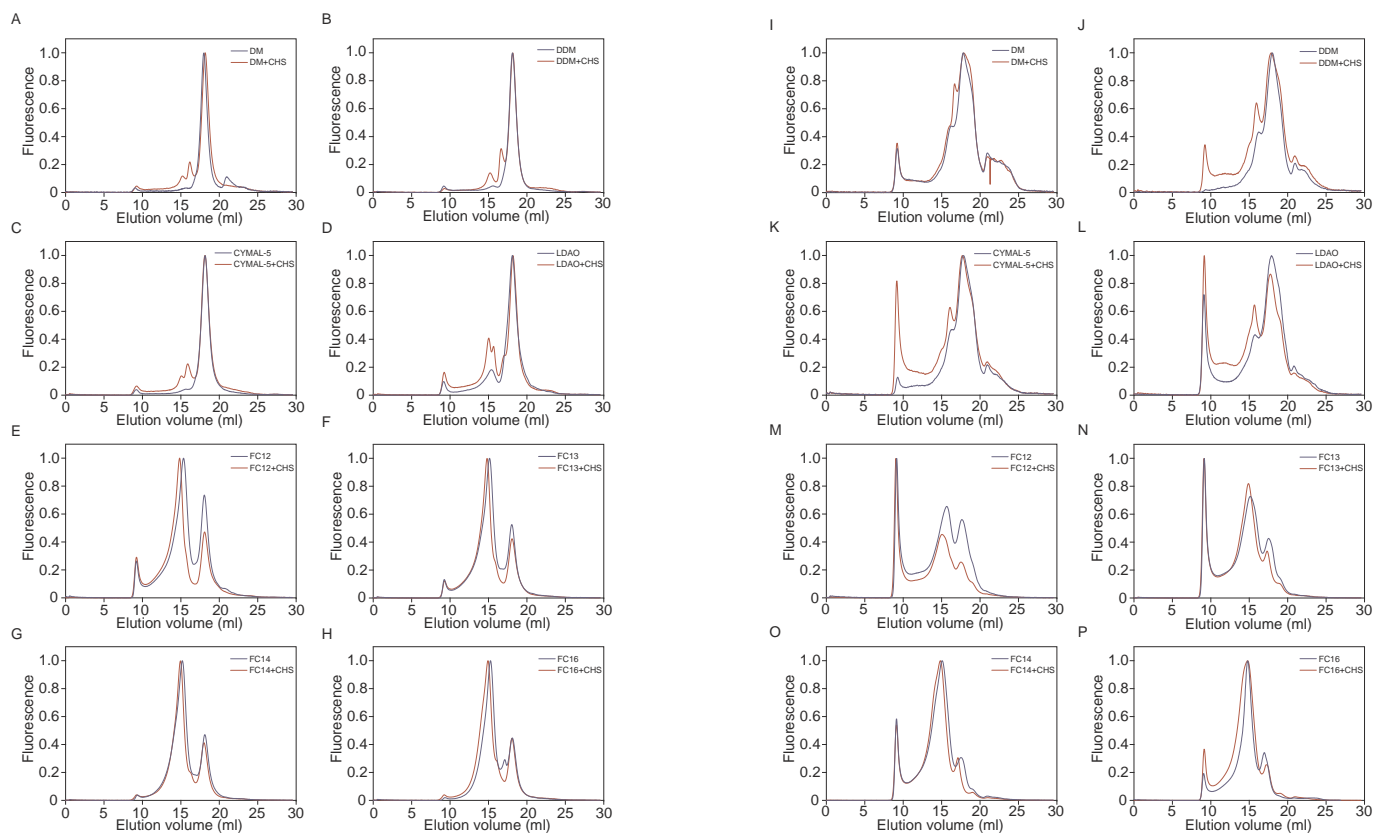


Fig. 6. Fluorescence-detection size exclusion chromatography (FSEC) of hNHE1 (A-H) and hGHR (I-P). FSEC was performed to evaluate the quality of solubilized hNHE1 and hGHR. Detergent solubilized hNHE1 and hGHR were run on a Superose 6 increase 10/300 GL column as described in **Materials and methods**. FSEC profiles of solubilized hNHE1 and hGHR in DM, DDM, Cymal-5, LDAO, FC-12, FC-13, FC-14 and FC-16 in the absence and presence of CHS. Fluorescence was normalized to the peak value in each profile. The void volume is approximately at 8 ml. The peak at approximately 18 ml is GFP liberated from the fusion protein.

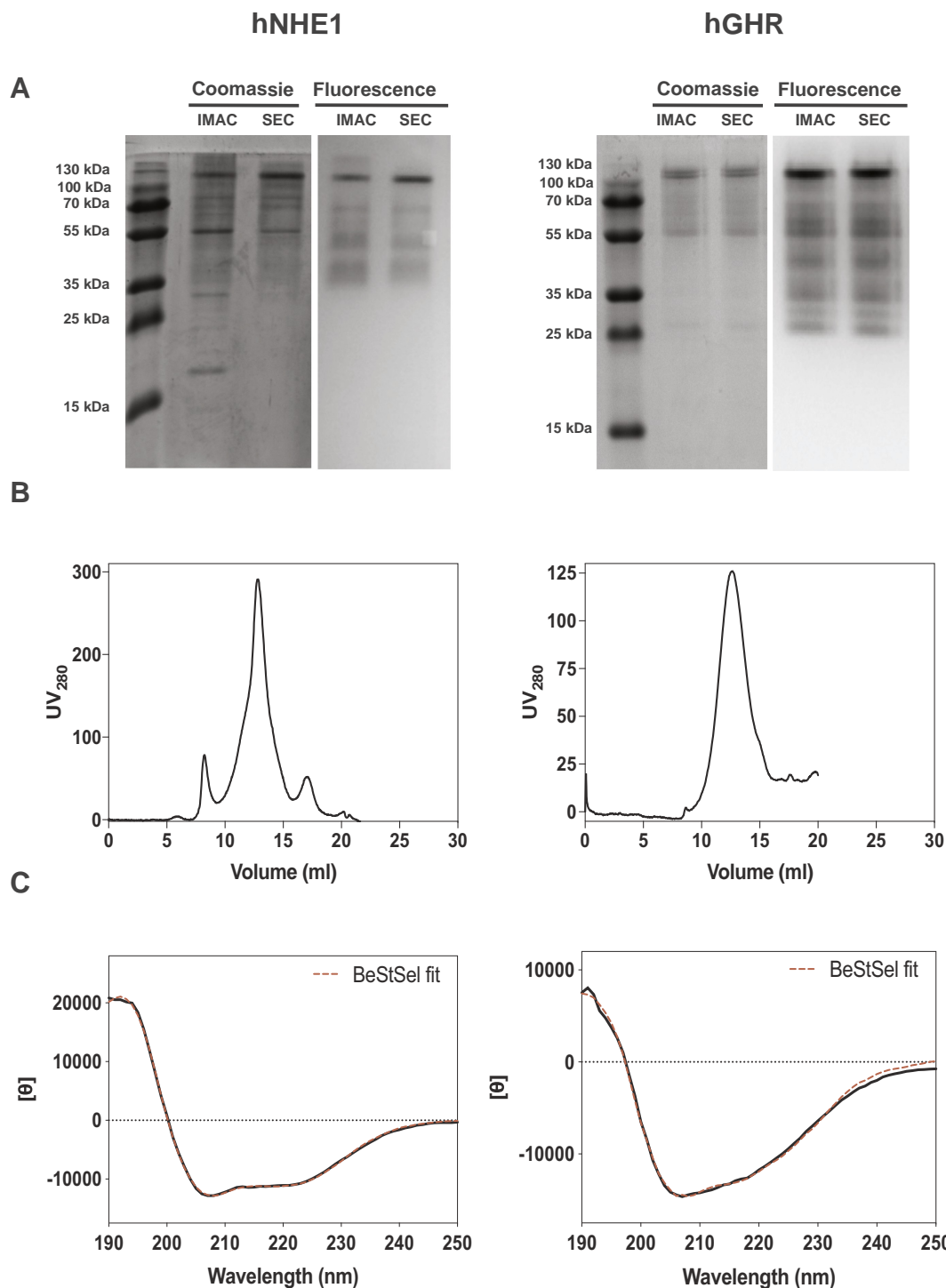


Fig. 7. Purification of hNHE1 and hGHR. (A) hNHE1 (left) and hGHR (right) were purified by Ni-affinity chromatography (lane: IMAC) and subsequently SEC (lane: SEC) and visualized by both Coomassie staining and in-gel fluorescence. (B) SEC analysis of Ni-affinity chromatography purified hNHE1 (left) and Ni-affinity chromatography purified hGHR (right). (C) Far-UV CD spectroscopy analysis of 1 μ M hNHE1 (left) hNHE1 and 1 μ M hGHR (right). Red dashed lines represent BeStSel analyses of secondary structure content. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Purity and yield for hNHE1 and hGHR per 1 L cell culture/12 g cells.

| Samples | IMAC yield from A ₂₈₀ (mg) | SEC yield from A ₂₈₀ (mg) | SEC yield from fluorescence (mg) | Purity |
|---------|---------------------------------------|--------------------------------------|----------------------------------|--------|
| hNHE1 | 1.1 | 0.70 | 0.59 | 84% |
| hGHR | 0.48 | 0.34 | 0.30 | 88% |

Table 3
Secondary structure content from model and from far-UV CD.

| | α -helix (%) | β -strand (%) | Turns and others (%) |
|------------------------|---------------------|---------------------|----------------------|
| hNHE1 _{CD} | 27 | 11 | 62 |
| hNHE1 _{model} | 23 | 9.0 | 68 |
| hGHR _{CD} | 11 | 29 | 60 |
| hGHR _{model} | 8 | 22 | 70 |

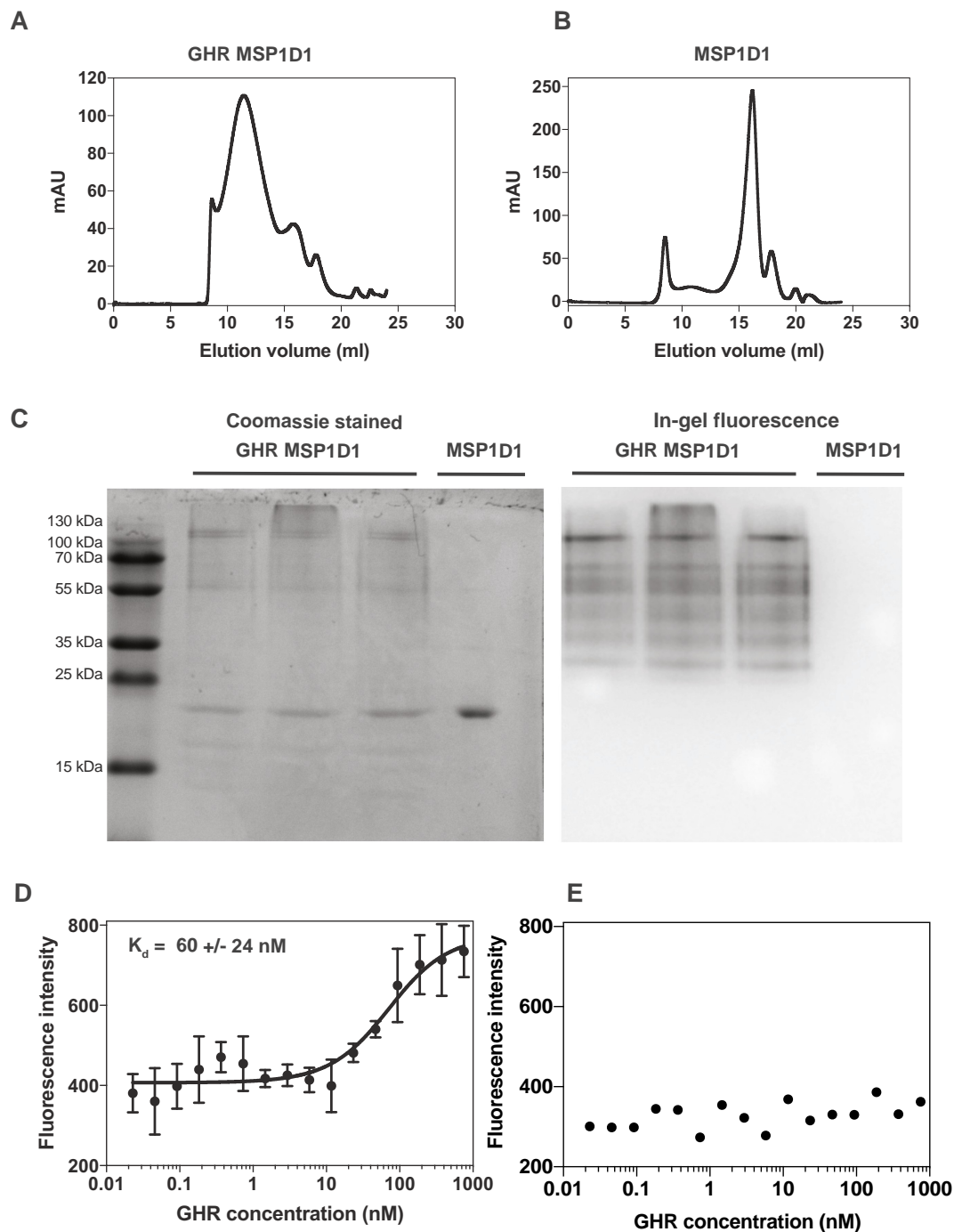


Fig. 8. Incorporation of hGHR into MSP1D1 nanodiscs. (A) SEC purification of hGHR loaded MSP1D1 nanodiscs. (B) SEC purification of empty MSP1D1 nanodiscs. (C) SDS-PAGE analysis of hGHR incorporated into MSP1D1 and empty MSP1D1 discs (D) MST of hGH-G120R binding to hGHR in a nanodisc. (E) MST of hPRL binding to hGHR in a nanodisc.

and constitute a system for which several different protocols have been published. Second, and more importantly, as membrane proteins with long IDRs are likely to be investigated by small angle scattering techniques, nanodiscs form more homogeneous samples preparations than e.g. SMALPS and saposin, and the small angle scattering parameters from the discs are readily obtained [71,72]. As hNHE1 forms a dimer with each monomer consisting of 12 TM helices [12], the most used nanodisc, the MSP1D1, is not a suitable scaffold and larger scaffolds would be needed. An extended MSP has been developed, the MSP1E3D1 [32], which should be adequate for 12 and 24 TM α -helices and this was used for testing reconstitution of hNHE1. For hGHR, with

its single pass architecture, we used the conventional MSP1D1 for reconstitution.

To incorporate hGHR in MSP1D1, we mixed purified hGHR with POPC and sodium cholate to form mixed micelles prior to adding purified MSP1D1 from *E. coli* and detergent-extracting Bio-Beads. After incubation, IMAC was used to separate the loaded hGHR MSP1D1 discs from the empty MSP1D1 discs. SEC purification of loaded and unloaded discs was performed separately on a Superose 6 increase 10/300 column (Fig. 8A,B). This revealed that hGHR in MSP1D1 elutes at 10–14 mL, which was confirmed by SDS-PAGE, indicating that reconstitution was successful (Fig. 8C). The empty MSP1D1 discs were run as

reference (Fig. 8B). We also tested the incorporation of hNHE1 into a larger nanodisc using MSP1E3D1. We initially tried reconstitution in this nanodisc using the same lipid solution as with hGHR, but hNHE1 aggregated upon detergent removal. We then attempted to use POPC:POPS mixtures, but the hNHE1 sample still aggregated upon detergent removal. To mimic the native membrane even further we attempted a mixture of soybean polar lipid extracts (Avanti, USA) and CHS. However, no incorporation was observed. Thus, we were not able to incorporate hNHE1 into nanodiscs.

To finally assess the functionality of the nanodisc-reconstituted GHR, we applied MST and the fluorescently labeled receptor antagonist, the hGH-G120R. We used the antagonist as opposed to an agonist, with the purpose of not introducing dimers across the discs and because the agonist at high concentration would lead to monomerization, complicating the fluorescent readout. With this experimental approach, the dissociation constant between hGH_{G120R} and hGHR(MSP1D1) was determined to be $K_D = 60 \pm 24$ nM (Fig. 8D). The affinities of hGH for hGHR-ECD have previously been reported as 1.2 nM and 3.5 nM for binding to the first and the second site of hGH, respectively, and the affinity for hGH_{G120R} is known to be highly similar or slightly lower [74,75]. As negative control, we used hPRL, which cannot activate hGHR *in vivo*, and indeed, we observed no binding (Fig. 8E). Thus, the nanodisc-reconstituted, yeast-produced intact full length hGHR is binding competent.

Functional analysis is the ultimate test of any purified protein. However, as hNHE1 exchanges one proton for one sodium ion, it is electrically neutral and the measurements of activity require reconstitution into liposomes and co-establishment of a proton-generating system, which is beyond the scope of the current work (see e.g. [76]). The correlation of the secondary structure content with an existing model, the homogeneity of the elution peak from SEC, as well as the low FC16 concentration used in the extraction, suggest that hNHE1 is correctly folded.

In conclusion, we were able to produce intact full length hGHR and incorporate it into MSP1D1 nanodiscs, which on a larger-scale production will enable future structural studies using either SAXS or SANS and potentially also cryo-EM. NHE1 was produced to homogeneity and was helical and with an intact disordered tail, yet incorporation of NHE1 in the largest available nanodisc, MSP1E3D1, was not achieved.

4. Discussion

Because of the lack of structural characterization of MPs containing intra- and extra cellular intrinsically disordered domains, this work aimed at providing a protocol for purification and incorporation of intact MPs with disordered regions into nanodiscs. This protocol successfully led to the purification of two highly different MPs, the hNHE1 and the hGHR. The protocol is straightforward and will help improve the ability to produce and study MPs also with long disordered tails by means of biochemical and biophysical techniques and thereby increase the rate by which MPs are structurally characterized. We employed the two widely different human MPs, NHE1 and hGHR and produced them in *S. cerevisiae*. hNHE1 was selected because no high-resolution structure of hNHE1 or any eukaryotic homologs exists. A crystal structure of the GHR-ECD [23], NMR structures of the GHR-TMD [55] and NMR characterization of the C-terminal disordered tail [29] are available, but no experimentally based full-length structure is available for GHR or any other cytokine receptor.

We produced hNHE1 and hGHR with a C-terminal GFP tag to simplify optimization of expression and purification and also exploited GFP as a folding reporter [61]. A potential disadvantage of using GFP fluorescence to quantify recombinant MP accumulation is that fluorescent degradation products contribute to the measured GFP fluorescence. hNHE1 produced in *S. cerevisiae* was mainly located in the plasma membrane. However, hGHR was mainly located in intracellular membranes as detected by *in vivo* fluorescence microscopy. Studies of

membrane proteins purified from internal membranes of *S. cerevisiae* have suggested that they are functional [36,40,42]. We also observed that both hNHE1 and hGHR produced at 15 °C accumulated with the expected molecular weights and that degradation was minor (Fig. 3) and that they maintained a folded structure when purified. Fluorescence, therefore, originates from accumulation of intact hNHE1 and hGHR.

Identifying suitable detergents for purification of MPs involves labor intensive screening, as a single detergent that fits all MPs does not exist. The detergent screen on hNHE1 and hGHR revealed that they are only soluble in zwitter-ionic detergents such as LDAO and FCs (Fig. 6), and not in the frequently used detergent DDM. High solubilization efficiency is of course a high priority, but it is even more important that the solubilization results in a homogenous MP sample. This was assayed by FSEC analysis of hNHE1 and hGHR using all tested detergents in the absence and presence of CHS (Fig. 7). Both hNHE1 and hGHR were homogenous in the FCs, and since FC-16 solubilized the most protein, this was used for purification (Fig. 7). Both hNHE1 and hGHR were isolated by Ni-affinity chromatography and subsequent SEC to high purity and homogeneity. Furthermore, both proteins remained folded as evaluated by CD spectroscopy.

Recent cryo-EM, SAXS and SANS studies of MPs have emphasized the importance of nanodiscs as a tool for *in vitro* studies of MPs under native-like conditions [69–72]. Indeed, even with a long disordered tail of > 350 residues, it was possible to incorporate hGHR into MSP1D1. Because of the size of hNHE1 - a 12 TM-helical MP that forms dimers in its native state - we attempted to incorporate hNHE1 into larger nanodiscs using MSP1E3D1. Unsuccessful attempts were made both using simple lipid solutions such as just POPC, but also in POPC:POPS mixtures and a mixture of soybean polar lipid extracts and CHS. As the C-terminal disordered tail of hNHE1 has lipid binding domains [12,77], the larger nanodisc may not have been large enough to allow solubilization *via* membrane interaction. It is therefore possible that a C-terminally truncated version of hNHE1 would be amenable for reconstitution into MSP1E3D1, but we did not pursue this further. A better solution which will allow studies of the intact NHE1 with its tail may be to use an even larger nanodisc, which is currently not available. Once these emerge, NHE1 would be a good first candidate for their application.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Mike J. Waters and Andrew J. Brooks for the generous gift of the hGH_{G120R} antagonist for MST studies and Lise Arleth, Nicolai Tidemand Johansen and Frederik Grønbaek Tidemand for providing the MSP1D1 plasmid and for a demonstration of how to reconstitute MPs into nanodiscs. Katrine Bugge is thanked for fruitful discussions and Signe A Sjørup for skilled technical assistance. This work was supported by The Novo Nordisk Foundation SYNERGY Grant (to S.F.P. and B.B.K) and - Challenge Grant, REPIN (B.B.K) and by the Lundbeck Foundation (to S.F.P and B.B.K). P.A.P and B.B.K are affiliated the *Integrative Structural Biology Cluster at UCPH – ISBUC*.

References

- [1] A. Krogh, B. Larsson, G. Von Heijne, E.L.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, *J. Mol. Biol.* 305 (3) (2001) 567–580, <https://doi.org/10.1006/jmbi.2000.4315>.
- [2] K. Lundstrom, Structural genomics and drug discovery, *J. Cell. Mol. Med.* 11 (2) (2005) 224–238, <https://doi.org/10.1111/j.1582-4934.2007.00028.x>.

- secondary structure prediction and fold recognition from the circular dichroism spectra, *Nucleic Acids Res.* 46 (W1) (2018) W315–W322, <https://doi.org/10.1093/nar/gky497>.
- [57] L. Counillon, J. Pouyssegur, R.A.F. Reithmeier, The Na⁺/H⁺ exchanger NHE-1 possesses N- and O-linked glycosylation restricted to the first N-terminal extracellular domain, *Biochemistry*. 33 (34) (1994) 10463–10469, <https://doi.org/10.1021/bi00200a030>.
- [58] F. Dehkhoda, C.M.M. Lee, J. Medina, A.J. Brooks, The Growth Hormone Receptor: Mechanism of Receptor Activation, Cell Signaling, and Physiological Aspects, *Front. Endocrinol. (Lausanne)* 9 (2018) 35, <https://doi.org/10.3389/fendo.2018.00035>.
- [59] P.A. Harding, X.Z. Wang, B. Kelder, S. Souza, S. Okada, J.J. Kopchick, In vitro mutagenesis of growth hormone receptor Asn-linked glycosylation sites, *Mol. Cell. Endocrinol.* 106 (1–2) (1994) 171–180, [https://doi.org/10.1016/0303-7207\(94\)90200-3](https://doi.org/10.1016/0303-7207(94)90200-3).
- [60] S.J. Frank, S.Y. Fuchs, Modulation of growth hormone receptor abundance and function: roles for the ubiquitin-proteasome system, *Biochim. Biophys. Acta* 1782 (12) (2008) 785–794, <https://doi.org/10.1016/j.bbadis.2008.06.001>.
- [61] E.R. Geertsma, M. Groeneveld, D.-J. Slotboom, B. Poolman, Quality control of overexpressed membrane proteins, *Proc. Natl. Acad. Sci. U. S. A.* 105 (15) (2008) 5722–5727, <https://doi.org/10.1073/pnas.0802190105>.
- [62] A.M. Seddon, P. Curnow, P.J. Booth, Membrane proteins, lipids and detergents: not just a soap opera, *Biochim. Biophys. Acta* 1666 (1–2) (2004) 105–117, <https://doi.org/10.1016/j.bbamem.2004.04.011>.
- [63] J.L. Parker, S. Newstead, Membrane protein crystallisation: current trends and future perspectives, *Adv. Exp. Med. Biol.* 922 (2016) 61–72, https://doi.org/10.1007/978-3-319-35072-1_5.
- [64] H. Ilgü, J.-M. Jeckelmann, M.S. Gachet, et al., Variation of the detergent-binding capacity and phospholipid content of membrane proteins when purified in different detergents, *Biophys. J.* 106 (8) (2014) 1660–1670, <https://doi.org/10.1016/j.bpj.2014.02.024>.
- [65] T. Kawate, E. Gouaux, Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins, *Structure*. 14 (4) (2006) 673–681, <https://doi.org/10.1016/j.str.2006.01.013>.
- [66] S. Contreras-Martos, H.H. Nguyen, P.N. Nguyen, et al., Quantification of intrinsically disordered proteins: a problem not fully appreciated, *Front. Mol. Biosci.* 5 (2018) 83, <https://doi.org/10.3389/fmolb.2018.00083>.
- [67] E. Szollosi, E. Háy, C. Szász, P. Tompa, Large systematic errors compromise quantitation of intrinsically unstructured proteins, *Anal. Biochem.* 360 (2) (2007) 321–323, <https://doi.org/10.1016/j.ab.2006.10.027>.
- [68] M. Landau, K. Herz, E. Padan, N. Ben-Tal, Model structure of the Na⁺/H⁺ exchanger 1 (NHE1): functional and clinical implications, *J. Biol. Chem.* 282 (52) (2007) 37854–37863, <https://doi.org/10.1074/jbc.M705460200>.
- [69] Y. Gao, E. Cao, D. Julius, Y. Cheng, TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action, *Nature*. 534 (7607) (2016) 347–351, <https://doi.org/10.1038/nature17964>.
- [70] H.E. Autzen, A.G. Myasnikov, M.G. Campbell, D. Asarnow, D. Julius, Y. Cheng, Structure of the human TRPM4 ion channel in a lipid nanodisc, *Science*. 359 (6372) (2018) 228–232, <https://doi.org/10.1126/science.aar4510>.
- [71] N.T. Johansen, F.G. Tidemand, T.T.T.N. Nguyen, K.D. Rand, M.C. Pedersen, L. Arleth, Circularized and solubility-enhanced MSPs facilitate simple and high-yield production of stable nanodiscs for studies of membrane proteins in solution, *FEBS J.* 286 (9) (2019) 1734–1751, <https://doi.org/10.1111/febs.14766>.
- [72] N.T. Johansen, M.C. Pedersen, L. Porcar, A. Martel, L. Arleth, Introducing SEC-SANS for studies of complex self-organized biological systems, *Acta Crystallogr Sect D, Struct Biol.* 74 (Pt 12) (2018) 1178–1191, <https://doi.org/10.1107/S2059798318007180>.
- [73] N. Skar-Gislinge, N.T. Johansen, R. Høiberg-Nielsen, L. Arleth, Comprehensive study of the self-assembly of phospholipid nanodiscs: what determines their shape and stoichiometry? *Langmuir*. 34 (42) (2018) 12569–12582, <https://doi.org/10.1021/acs.langmuir.8b01503>.
- [74] B. Bernat, M. Sun, M. Dwyer, M. Feldkamp, A.A. Kossiakoff, Dissecting the binding energy epitope of a high-affinity variant of human growth hormone: cooperative and additive effects from combining mutations from independently selected phage display mutagenesis libraries, *Biochemistry*. 43 (20) (2004) 6076–6084, <https://doi.org/10.1021/bi036069b>.
- [75] J.-L.K. Kouadio, J.R. Horn, G. Pal, A.A. Kossiakoff, Shotgun alanine scanning shows that growth hormone can bind productively to its receptor through a drastically minimized interface, *J. Biol. Chem.* 280 (27) (2005) 25524–25532, <https://doi.org/10.1074/jbc.M502167200>.
- [76] C. Lee, H.J. Kang, C. Von Ballmoos, et al., A two-domain elevator mechanism for sodium/proton antiport, *Nature*. 501 (7468) (2013) 573–577, <https://doi.org/10.1038/nature12484>.
- [77] N. Shimada-Shimizu, T. Hisamitsu, T.Y. Nakamura, N. Hirayama, S. Wakabayashi, Na⁺/H⁺ exchanger 1 is regulated via its lipid-interacting domain, which functions as a molecular switch: a pharmacological approach using indolocarbazole compounds, *Mol. Pharmacol.* 85 (1) (2014) 18–28, <https://doi.org/10.1124/mol.113.089268>.