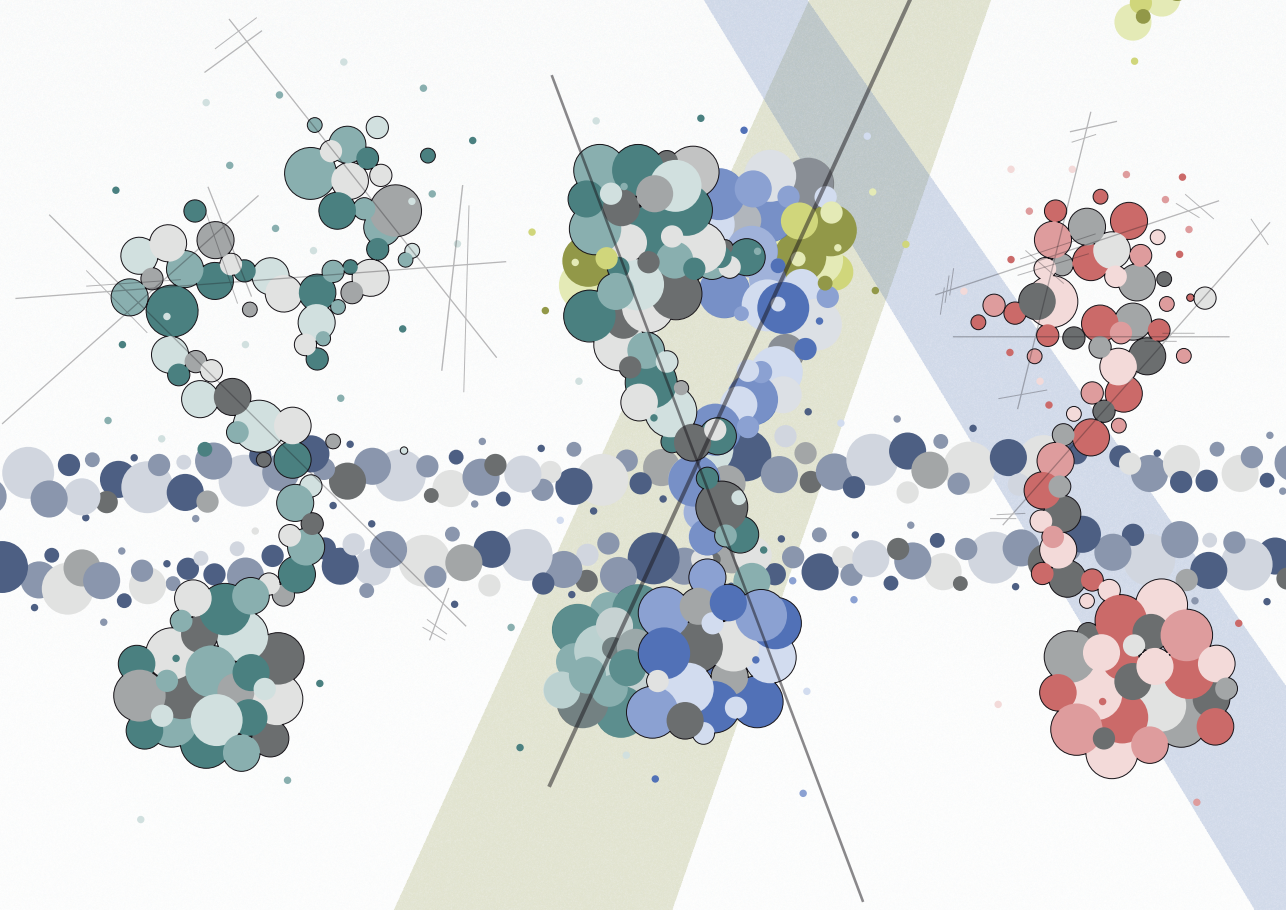
On the dynamic relationship between protein structure and function in native environments

Reinier Damman



On the dynamic relationship between protein structure and function in native environments

ISBN: 978-90-393-7278-4

Doctoral Thesis
On the dynamic relationship between protein structure and function in native environments

The studies described in this thesis were performed at the division of NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht University, The Netherlands and the division of Cell Biology, Neurobiology and Biophysics, Department of Biology, Faculty of Science, Utrecht University, Utrecht, the Netherlands.

Cover: Molecules in motion

Cover design by Cátia P. Frias and Reinier Damman.
Layout by Cátia P. Frias and Reinier Damman.
Printed by ProefschriftMaken.

Copyright © 2020 Reinier Damman
All rights reserved.

On the dynamic relationship between protein structure and function in native environments

Over de dynamische verhouding tussen eiwit structuur en functie in natieve omgevingen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op maandag 20 april 2020 des middags te 4.15 uur

door

Reinier Damman

geboren op 4 oktober 1992 te Zwolle, Nederland

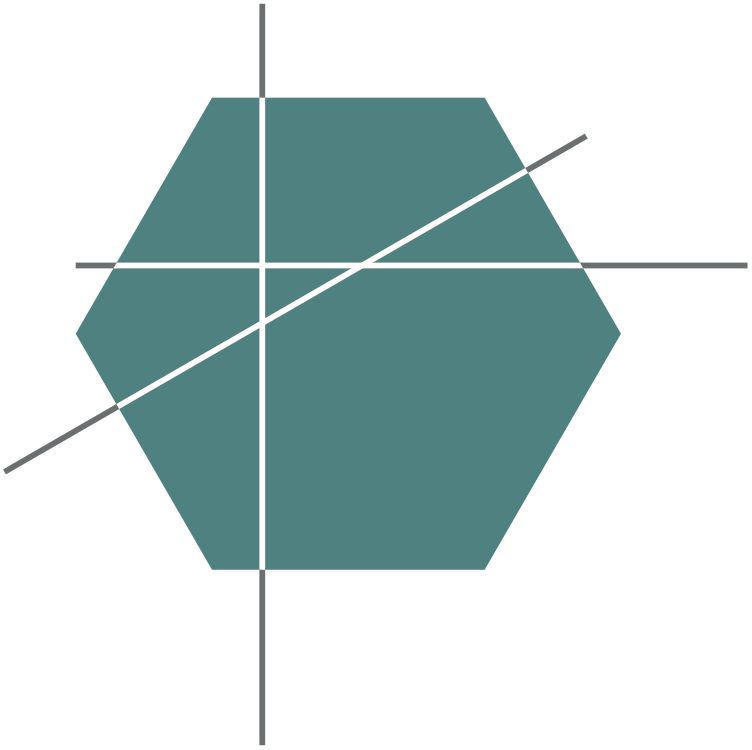
Promotor: Prof. dr. M. Baldus

Copromotor: Dr. P. M. P. van Bergen en Henegouwen

Aan mijn ouders

TABLE OF CONTENTS

Chapter 1	General Introduction	9
Chapter 2	Cellular Solid-state NMR Spectroscopy	33
Chapter 3	Allosteric nanobodies reveal the importance of regulating conformational dynamics in EGFR functioning	53
Chapter 4	Rigidification of the extracellular domain of EGFR stabilizes HER2 and stimulates its functioning	73
Chapter 5	Atomic-level insight into mRNA processing bodies by combining solid and solution-state NMR spectroscopy	89
Chapter 6	Development of in vitro-grown spheroids as a tumor model system for ssNMR	121
Chapter 7	General Discussion & Future Perspectives	137
Addendum	Summary	146
	Samenvatting	148
	Curriculum vitae	150
	List of publications	151
	Acknowledgments	152



Chapter 1

General Introduction

Reinier Damman

NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands

Cells rely heavily on their potential to change rapidly and accordingly to a continuously changing milieu. They utilize surface-exposed proteins, receptors, which continuously sense the cellular environment for outside signals. When a receptor does interact with a molecule, it will trigger a response inside the cell that induces changes to the proteome, thereby affecting cell viability and functioning. Understanding how proteins behave at atomic resolution and function inside their native environment is of vital importance to comprehend the molecular processes that drive cells in health and disease.

For many years, protein structures have been solved by X-ray Crystallography, solution-state NMR and Cryo-Electron microscopy. However, each of these techniques has inherent disadvantages which limits the translation of structural data towards models of activation mechanisms that also include information obtained from molecular biology, as the experimental conditions differ tremendously. This is especially evident for membrane-embedded proteins. Here solid-state NMR (ssNMR) presents a unique quality as it is able to generate data at atomic-level resolution even in the context of large complex systems. Moreover, it also can record both motion- and rigidity-sensitive experiments allowing for a full characterization of the protein of interest.

In this thesis we attempt to elucidate the complex relationship between protein structure and function while studying them under physiological conditions. We also developed a novel 3-dimensional cellular sample methodology for ssNMR purposes that allows direct studies on tumor model systems. Our work indicates the necessity for domain motion in processes like growth factor receptor activation and liquid-liquid phase separation. The following section with describes the fundamentals on the structure-function relationship of proteins and the importance of the native environment. Finally, the fundamentals of ssNMR are discussed.

STRUCTURE-FUNCTION RELATIONSHIP

For many decades, structural biology has given insights into the atomic structure of proteins and protein complexes. The knowledge gained from these studies has vitally contributed to a better understanding of how proteins function and drive cellular homeostasis. Indeed, advancements in the field of structural and molecular biology concomitantly contributed to a better comprehension how protein structure relates to function. Many receptors structures have been solved in both the apo- or ligand-bound state, thereby elucidating the changes that occur after ligand binding ranging from small rearrangements of amino-acid side chains or loop regions to large conformational changes where individual subdomains move significantly in respect to one another.

As the apo- and bound-state usually exhibit different conformations, molecular motion is necessary to trigger conformational changes upon ligand binding necessary for function. Moreover, conformational heterogeneity would allow for a certain flexibility and heterogeneity that directly influences the regulation of vital parameters e.g. activity, ligand binding affinity or oligomerization. For example, allosteric inhibitors can enforce a "closed" or "tethered" state, keeping NMDA receptors in an inactive conformation, distinctly different from several structures of the apo-state[1–3]. In addition, extensive research performed on G-protein coupled receptors (GPCR) highlighted the existence of activation of distinct

signaling pathways depending on the binding of ligands to the receptor. The orientation of a single transmembrane helix controls the binding affinity of a ligand, thereby allowing for tight regulation of the activation of downstream signaling pathways[4,5]. For receptors that rely on (hetero)dimerization for their signal transduction, the binding of different ligands will lead to slightly altered structures that drive (hetero)dimerization and activation in distinct fashion[6–8].

Epidermal growth factor receptor and Her2

The epidermal growth factor receptor (EGFR) is a member of the human EGFR family of receptor tyrosine kinases (also known as HER family or ErbB family) together with Her2, Her3 and Her4. Extensive studies on these receptor tyrosine kinases by structural- and molecular biology have led to a general understanding of the complex functioning of these receptors after stimulation by a variety of activating ligands. Crystal structures, complexed to different ligands, showcase the diverse conformations that potentially occur on the plasma membrane which drive intracellular signing transduction pathways that initiate cellular proliferation, differentiation, migration and inhibition of apoptosis. Despite playing a crucial role in homeostasis of tissue, irregularities in expression, activation and down-regulation of EGFR family members are often implicated in tumorigenesis[9–13].

The general architecture of the EGFR family is conserved. The ligand-sensing extracellular domain (ECD) consist of 4 subdomains, usually referrers to as domain one to four (DI to DIV), which is connected via a single transmembrane spanning helix domain (TMD) and juxtamembrane domain (JMD) to the intracellular kinase domain (KD). A long unstructured C-terminal tail (CT) contains several residues that become phosphorylated upon receptor activation (Fig 1.b). The first crystal structures of the EGFR ECD revealed that the ligand-binding domains DI and DIII share a similar structure, while the dimerization domains DII and DIV share similar features[14–16.] Soluble ECD (sECD) adopts a closed or "tethered" conformation where the β-hairpin of DII rests upon DIV (Fig 1.b). However, it is thought that resting EGFRs in the plasma membrane of cells are present in several oligomeric states[17–21], which play a role in the autoinhibitory regulation of EGFR. The large majority of EGFR is monomeric, while smaller population of dimer, trimeric or even higher oligomeric states are found[22,23]. Three different oligomeric structures have been described so far: i) a ligand-free assembly where monomers stack side-by-side[15,24,25], ii) an "extended"-but ligand-free back-to-back dimer, resembling the low-affinity dimer interface as seen for Drosophila ECD[26,27] and iii) a ligand-free stalk-to-stalk interaction mediated by DIV and the TMD[28]. The later model proposes that DI-DIII of the EGFR ECD are still mobile[28,29], which is in line with ssNMR data in native plasma membrane vesicles[30]. Recently, a fourth autoinhibitory structure has been proposed by Zanetti and colleagues, in which the EGFR ECDs form a polymer chain, mediated by trans interactions between DI and DII but also by interactions between DIII and DIV[23]. In this structure, the intracellular KDs are well-separated in space for phosphorylation to occur, thereby autoinhibiting EGFR functioning[23]. However, they propose that the polymer structure stimulates a rapid response to growth factor stimulation[23]. It is likely that these structures co-exist in equilibrium on the plasma membrane, changing between oligomeric states and structures, thereby potentially directly influence receptor availability, binding affinity and activation rate.

One of the main functions of the ECD is to bind activating ligands. Up to this point, a total of 12 activating ligands have been identified for EGFR including the epidermal

growth factor (EGF), transforming growth factor α (TGF-α), amphiregulin (AREG) and epiregulin (EREG)[31,32]. Crystal structures of sECD in complex to any of these ligands reveal that the ECD adopts a significantly different structure after ligand binding. As DI and DIII are simultaneously involved in ligand binding, a crucial pivot point between DII and DIII allows for a rotation of 130°, exposing the dimerization arm of DII and dimerization motifs on DIV[14,33]. In native plasma membranes, it has been shown that monomeric EGFR is highly dynamic and continuously samples the conformational landscape before EGF stabilizes the ECD in the extended conformation[30]. Depending on the activating ligand, EGFR ECDs adopt slightly different conformations, suggesting a structural heterogeneity[34]. The Her2 ECD resembles that of the EGFR, however, the Her2 ECD has an additional loop in DI which occupies the space which is occupied by e.g. EGF for EGFR[35,36]. Interestingly, due to this additional loop, DI and DIII are directly interacting with one another, and consequently, the conformation in which the Her2 ECD is crystallized globally resembles the conformation of ligand-bound EGFR ECDs, suggesting a constitutively open conformation for Her2[35,37]. There has no activating-ligands found that directly binds to Her2 and induce phosphorylation, likely due to the additional loop in DI[36,38].

Exposure of the DII- and DIV- dimerization motifs in EGFR- and Her2 ECD, the ECDs (hetero)dimerize, bringing the TMD, JMD and KD in close proximity of each other. Dimerization is further supported by motifs in the TMD and JMD that tightly interact close to the membrane, maintaining the kinase domain in an inactive state[39–42]. Oncogenic

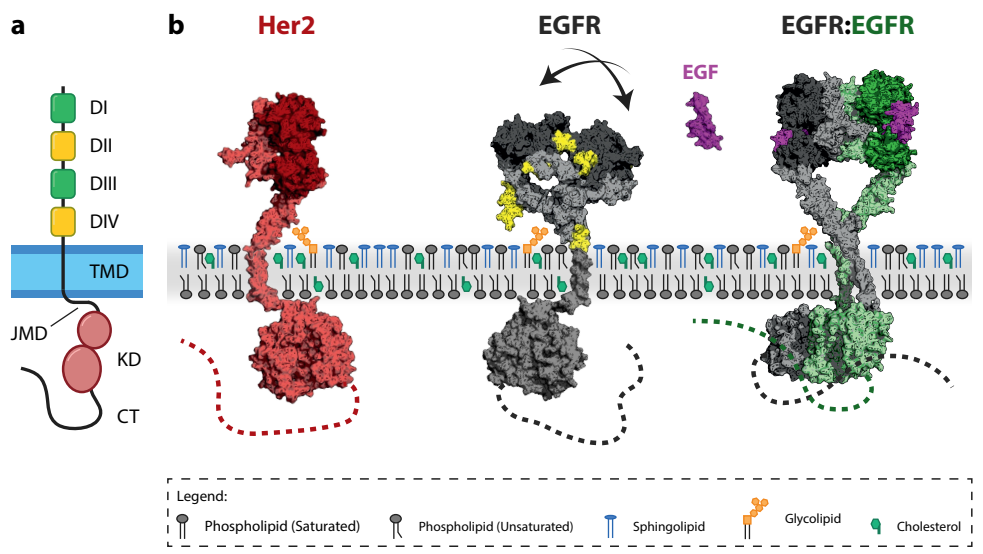


Figure 1 | Atomistic models of Her2 and EGFR monomers and EGFR dimer. a, general architecture of the EGFR tyrosine kinase receptor family. The ECD consists of 4 subdomains of which DI and DIII (green) are involved in ligand binding. DII and DIV play a role in dimerization (yellow). The ECD is connected to the intracellular domains via a transmembrane domain (TMD) and juxtamembrane domain (JMD). The kinase domain (KD) is highlighted in red followed by a long unstructured C-terminal tail (CT). **b,** Atomistic monomer models of Her2 (red; PDB: 1N8Z, 2JWA, 2KS1, 3PP0) and EGFR (gray; PDB: 1NQL, 1M14, 1Z9I, 4KRO) and an EGF-induced (magenta; PDB) EGFR homodimer (gray and green; PDB: 1IVO, 1MOX) in a native plasma membrane. N-glycan modifications on EGFR are highlighted in yellow (adapted from Kaszuba et al. 2015).

mutations in the JMD of Her2 release the JMD from the negatively charged lipids leading to increased dimerization, highlighting the importance of oligomerization on tumorigenesis[43–45]. Further investigations showcase the complex structural heterogeneity in receptor activation has functional implications, as JMD interactions in a dimer depend on ligand binding to the ECD[46,47].

The highly conserved tyrosine kinase domain joins 2 globular domains, the smaller N-lobe and the C-lobe. ATP binds tightly within a cleft between the N- and C-lobe with a conserved glycine-rich loop in the N-lobe. The N-lobe also hosts the active site for the catalysis of ATP. The C-lobe supports the N-lobe by a conserved catalytic loop (D812-N818). Moreover, the C-lobe regulates the availability of the active site by means of an activation loop[48–51]. In asymmetric kinase dimers, the C-lobe of the "activator" binds close to the N-lobe of the "receiver", which in turn shifts the activation loop and opens up the active site. The KD is commonly mutated in cancer, which often leads to increased EGFR activity. For example, certain point mutations abolish the autoinhibitory effect of the activation loop, while others shift the oligomeric state of EGFR form monomers to higher order clusters[52–54]. Finally, signal transduction is initiated as the C-termini get transphosphorylated on specific residues, thereby recruiting adaptor proteins to transfer the signal. It is thought that the large majority of the CT is unstructured. Even so, it is proposed that the CT plays an important regulatory role as the KD cycles through active- and inactive states. Conformational changes in the CT, but also the JMD, support the ATP binding and kinase dimerization[55,56].

Receptor dimerization is required for its activation, but also plays an important role in down-regulation of actively signaling complexes[39]. Additionally, higher order oligomeric complexes have been proposed for EGFR and Her2[57–59]. The Kuryian lab proposed an oligomeric structure, where EGF-bound EGFR dimers interact side-by-side via DIV to form tetramers and potentially propagate to form even higher order oligomers[60]. In this model, every EGFR molecule is bound to EGF. Other models suggest that in EGFR oligomers, the outside receptors are only bound by EGF, leaving available binding epitopes between DI and DIII for unliganded receptors to interact via front-to-front dimers[61]. Interestingly, they show that the distribution of EGFR oligomeric state is dependent on the EGF concentration, where high EGF concentrations drive EGFR dimerization[61].

Activated EGFR complexes actively trigger intracellular signalling cascades, but will also induce downregulation pathways. There are two major downregulation pathways described for EGFR. Generally, EGFR is internalized via Clathrin-mediated endocytosis (CME), after which endosomes are majorly sorted to be recycled, containing inactivated EGFRs, back to the plasma membrane or they are trafficked to lysosomes for degradation[62.] However, elevated EGFR expression levels or an increase in local EGF concentration, can lead to significant increase in ubiquitination of EGFR, which triggers the Clathrin-independent endocytosis (CIE) pathway, leading to subsequent degradation of EGFR[63–65]. Therefore, it is thought that CIE is the default pathway for EGFR degradation, while CME is mainly involved in recycling of inactivated receptors[66]. The section above described the presence of dimerization motifs on DII and DIV on the EGFR ECD. Additionally, two dimerization motifs have also been found in the TMD. The N-terminal GxxxG motif is crucial in clustering-induced internalization of EGFR and could be site-specifically crosslinked to each other after dimerization[67,68]. On the other hand, the C-terminal dimerization motif is not involved in CME and potentially has a role in different EGFR functions. Upon activation of EGFR, different lipids are recruited to form nanodomains in the plasma membrane[69,70] (Also see

section below). These lipids may act to stabilize the N-terminal dimer interaction of the TMD and allow EGFR clusters to bend the plasma membrane[61,71]. This will locally lower the lipid pressure and results in the recruitment F-BAR proteins, amongst other BAR domain containing proteins, that stabilize the concave membrane structure, promoting the initiation of receptor internalization[72–74]. Due to the close relationship between dimer (or oligomer) structures and signalling, it is likely that this structural heterogeneity will also influence the sorting, recycling and degradation of EGFR.

Taken together, the allosteric coupling of the ECD with the intracellular domains and their role in receptor activation and signaling highlight the complex behavior of receptor functioning. Individual ligands stabilize a variety of structures that all have a different effect on signaling and down-regulation[34]. Moreover, it has been proposed that motion of the ECD is crucial for ligand binding and subsequent activation[30]. This is further supported by point mutations found in the ECD of EGFR of glioblastoma patients which link increased binding affinity of ligands to increased ECD mobility[29,75]. Additionally, increased conformational freedom allows for the stabilization of complexes that drive (hetero)dimerization. Yet, a role for the ECD dynamics in the activation of these receptors has never been proposed.

RNA-proteins

Many cellular processes occur in specialized organelles, which are isolated from one another by membranes. However, renewed interest in membrane-less organelles has emerged in the last decade leading to the idea that protein compartmentalization is more wide-spread phenomenon than initially thought and can have major implications on cell biology. Membrane-less organelle formation is related to the process of a liquid-liquid phase separation (LLPS) in which proteins can be present in different phases depending on the chemical environment, ranging from a soluble well-dispersed solute phase, that transit via a condensed- and gel-like phase into a solid- or fibril phase (Fig 2). It is believed that the multivalent weak interactions jointly contribute to the condensation of proteins (and possibly other biomolecules such as RNA) into these condensed states. As interactions between proteins mature during LLPS, interactions strengthen and could result in conformational changes, leading to a solid-phase[76]. Interestingly, it has been proposed that the dynamic behavior of these proteins is important for their functioning and can differ between proteins. For example, Balbiani bodies in oocytes resemble a solid-phase while the naturally more dynamic RNA-binding proteins (RNP) remain liquid-like after LLPS[77–79]. However, when maturation of LLPS is not regulated, even dynamic protein assemblies tent to aggregate, which has been implicated in neurodegenerative diseases such as ALS[80,81].

In short, as a consequence of phase separation, proteins sequester can cluster at a high local concentration to perform or facilitate a range of biological processes. Firstly identified in the nucleolus, LLPS aids in the nucleation of for example ribosomes or the cytoskeletal network[82–84]. Moreover, the condensed state of proteins supports catalytic processes and signal transduction[84,85]. Removing proteins from the active pool might also inactivate or block certain processes as catalytically active proteins are not able to interact with their substrates. A secondary role can be imagined as sequestered proteins can also recruit other proteins into the condensed state, thereby indirectly supporting catalytic activity[86–88]. Thirdly, proteins undergo LLPS as a fast and temporary response to external stimuli e.g. stress. In pH- and temperature sensing, changes in the milieu of cells are detected and induce phase separation as a means to rapidly alter their functional proteome, thereby

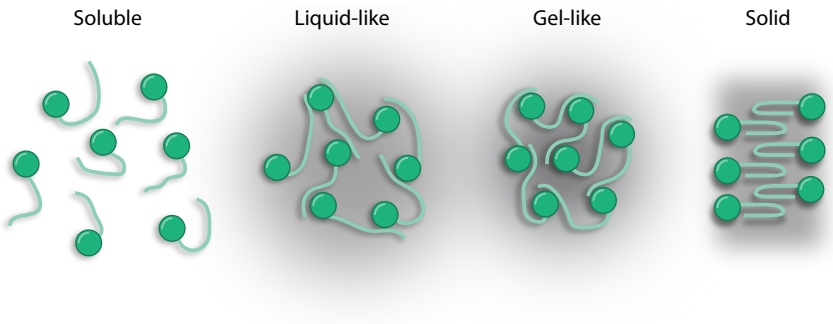


Figure 2 | Schematic representation of different phase-separated states. The dynamic state of proteins differs depending on their macromolecular organization, ranging from fully dynamic when they are soluble to solid- or fibril state. Weak intermolecular interactions allow for the (transient) occurrence of liquid- and gel-like states.

increasing its viability[89–91]. This response is much faster than the induction of transcription and translation. Phase separation potentially stores these proteins to later be released into the cytosol again. It has been proposed that a similar behavior is to be expected when protein concentrations exceed physiological concentrations[92]. Finally, the condensed state of proteins after phase separation has been proposed to be involved in forming a physical- and/or chemical barrier that filters the passage of molecules, thereby regulating passage of molecules and initiating polarization in cells[93].

Phase separated proteins appear as foci with a spherical morphology on mesoscale level in vitro and in vivo, closely resembling oil-in-water structures, therefore often termed liquid droplets. The exchange rate of phase-separated proteins is in the order of seconds, rather than minutes or even longer as frequently observed for stable complexes[94,95]. Additionally, liquid droplets only have a spherical morphology in resting state, but this globular structure is not fixed and can easily change its shape when encountering another object or when cytoplasmic movements force liquid droplets to change their position[96]. Moreover, liquid droplets can fuse together or split into two separate droplets[96]. The dynamic behavior of these complexes is important as the liquid droplet state is only adopted temporarily and should dissolve again after a number of time.

Due to the dynamic behavior of the liquid droplets, structural and dynamical information at an atomic level have been limited thus far. However, three determinants of protein-protein interactions have been identified that give rise to liquid droplet formation: 1) well-folded subdomains that interact with other oligomerization domains to form multivalent network, 2) dimerization domains that are connected via flexible linkers to form a linear oligomerization strand, or 3) intrinsically disordered regions (IDR) of that form an interaction site for other subdomains or biomolecules[97,98]. The amino acid sequence of the IDR regions is of low complexity and very repetitive, mainly consisting of QGSY- or G-rich motifs[99,100]. Many studies have been performed on complexes of low-complexity IDRs by crystallography and NMR which all show a similar trend, where strong interactions are formed between β-strands to form a fibrillary structure[101,102]. The low complexity region in FUS, a protein that has been shown to undergo phase transition, also forms fibrils as seen by cryoEM and ssNMR[103]. In contrary, the structure of truncated subdomains of full-length proteins have been solved by solution NMR or crystallography from which models

have been derived of how these proteins behave in solution. However, this leaves a gap between the liquid- and solid-phase for which there are only structural models available. As a result, insight into whether and which conformational changes occur when going from the liquid- to a condensed state and how maturation of the condensed state progresses into protein fibrils is currently limited. The first publications with structural information on phase transition used the power of solution- or solid-state NMR to study both rigid and mobile components of the proteins before and after phase separation, suggesting that dynamics plays an important role in the regulation of LLPS[104–106].

NATIVE ENVIRONMENT

Even though extensive studies on the molecular structure and dynamics by NMR, X-ray crystallography and EM have contributed to a better understanding of proteins in health and disease, a direct translation of protein structure to functioning has been limited due to the sample requirements for high-resolution structure determination. Proteins are often purified and reconstituted in simplified systems. Domain motion usually reduces the quality of the structural data, which means that they are normally stabilized by additives or truncated to single protein domains. Moreover, membrane proteins are even more challenging as they are embedded in or anchored to a dynamic lipid bilayer. Additionally, many eukaryotic proteins are post-translationally modified (PTM), e.g. glycosylated, which modulate protein function. Often, these PTMs are removed before attempts are made to solve the structure of the protein. Below is an overview of how these factors influence protein structure and function.

The plasma membrane

The plasma membrane forms a barrier between the cytosol and its extracellular milieu which is essential for the cellular homeostasis. The non-covalent interactions of lipids and membrane proteins maintain a dynamic and fluid structure in which they can freely move. Cells rely on a vast diversity of membrane proteins for their viability, ranging from cellular adhesion, chemosensing, ion- and metabolite transport and migration. It is estimated that around 30% of the mammalian genome encode for membrane proteins.

Lipids in the bilayer have a general architecture consisting of a hydrophilic headgroup connected to long hydrophobic tails. Cells regulate the membrane fluidity and thickness by adjusting the membrane composition for example by incorporating lipids with different tail lengths or introduce double-bonds in the tails to disturb hydrophobic interactions within the bilayer. The mammalian cell membranes main constituents are phosphatidylethanolamine (PE), phosphatidylserine (PS) and phosphatidylcholine (PC). Additionally to the phospholipids, eukaryotic plasma membranes also contain cholesterol and glycolipids. The molecular composition also differs between the outer and inner leaflet. The structure of cholesterol differs from that of phospholipids, where a rigid steroid ring is connected to a short hydrophobic carbon tail that can incorporate itself into the tails of phospholipids. Glycolipids are almost exclusively found on the outer leaflet of the plasma membrane (Fig 1.b). The most complex family of glycolipids, the ganglioside, contain at least one sialic acid modification which gives the ganglioside a net negative charge. The large variety of lipids and there unique attributes also results in the formation of lipids rafts, where local increases in lipid concentrations compartmentalize the plasma membrane. This also has

functional implications as membrane proteins also have preferred surrounding lipids, thereby creating complex mosaic protein distribution.

The importance of the native lipid membrane in structural biology is illustrated by studies that clearly show structural differences between membrane mimetics and native plasma membranes. The overall protein structure and oligomerization in native membranes can differ which has consequences on substrate binding and functioning[107,108]. Even for small antibiotics, its structure and binding on target cells differs from reconstituted bilayers[109]. These studies support the idea for structural studies on membrane proteins to be performed under conditions in the native environment.

The lipid environment also has big implications for the functioning of EGFR after EGF binding[110]. As EGFR favors a lipid composition rich in gangliosides and cholesterol, their presence also regulate EGFR[111]. Removal of cholesterol from the plasma membrane stimulated EGFR activity[21,112]. Moreover, cholesterol reduces the mobility of EGFR[113]. In contrast, GM3 (monosialodihexosylganglioside) negatively modulates kinase activity of EGFR via its N-linked glycans[114], potentially by reducing receptor dimerization[115–117]. It is believed that GM3 interacts with DIV or the EGFR ECD thereby preventing EGF to be able to induce phosphorylation of the receptor[118]. Ganglioside-dependent EGFR inhibition is also indirectly modulated as GM3 redistributes caveolin-1, which is thought to support activated receptors complexes and their respective adaptor proteins[117]. Moreover, EGF-induced activation of EGFR changes the local lipid composition of EGFR, which now co-localizes with GM1 (monosialotetrahexosylganglioside) that could support either dimerization of downstream signalling[119]. The importance of lipid composition in EGFR functioning is further highlighted by the upregulation of ganglioside synthase enzymes which support the activation and signaling of growth factor receptors[120,121]. Taken together, biological evidence suggests an important role for the lipid membrane in the regulation of EGFR structure and function. This is further underlined by studies performed on the full-length EGFR in native mammalian cancer cell line derived plasma membrane vesicles, which clearly illustrate that the ECD of EGFR exhibits high conformational freedom and sampling the conformational landscape prior to binding of EGF, which subsequently stabilizes the ECD to drive dimerization[30].

Glycosylation

Glycosylation is one of the types of post-translational modifications that can occur on proteins. Even though the mammalian glycome is thought to be larger than the proteome, only a smaller subset of glycan structures are found due to constrains and regulation on the glycosylation process. Glycans form covalent chains of polysaccharides to specific residues, thereby contributing to function. They have been thought to play a role in adhesion, trafficking, but more interestingly also in receptor activation, signal transduction and endocytosis[122].

As newly synthesized receptors are transported from the ER via the Golgi apparatus to the plasma membrane, where glycosyltransferase- and glycosidase enzymes modify specific residues, contributing to complex glycan linkages. The structure of the glycan linkages is cell-type specific and can be easily regulated by changes in the transcription of these enzymes and play a role in tumorigenesis[122,123]. Glycans impose a (allosteric) structural modification that can be directly involved in function. Due to the structural rearrangement, protein-protein interactions are promoted as new binding epitopes appear or they can be inhibited due to physical hindrance of dimerization interfaces[124–126]. For example, a neuronal adhesion molecule NCAM relies on its negatively-charged glycan moiety that is crucial for

NCAM to form synapses during development[127,128].

Glycosylation of receptors has distinct effects on its activation mechanism. The addition of sugar moieties forces conformational changes to the receptor that potentially shifts the equilibrium of oligomeric states on resting cells or exposes dimerization domains to allow for rapid activation[129,130]. Additionally, glycan modifications on lipids can also regulate receptor availability, clustering and diffusion- and internalization rates[131–133].

The ECD EGFR is also heavily glycosylated with approximately 25% of is molecular weight originating from the N-glycan moieties[134,135]. In total 9 out of 10 Asn residues in the ECD are glycosylated, giving rise to a diverse set of 55 glycan structures[136] with different roles in the regulation of EGFR[137]. It is thought that the presence of the glycans lifts the ECD, especially DI and DIII, of EGFR from the plasma membrane, potentially increasing receptor availability for binding[138]. Glycosylations found in DIV strengthen the tether of DII and DIV possibly leading to a reduction in activation. Other studies showed by molecular dynamics simulations on EGFR revealed that the glycan moieties also contribute EGF binding to dimer stability[139,140]. The importance is further highlighted by the changes in glycosylation patterns that have been found in cancer patients[123]. Upregulation of the *Mgat5* gene by the Ras-Raf-MAPK pathway results in an increase of branched N-linked glycans, which form a substrate for galectins. On its turn, galectins bind to branched glycan moieties forming galectin-glycoprotein, like EGFR, complexes on the membrane and reduce receptor internalization[141,142].

Taken together, glycosylation of the EGFR ECD has strong structural implications that are directly linked to functional aspects including ligand binding, binding affinity and dimer stability as well as disease.

NUCLEAR MAGNETIC RESONANCE

Nuclear magnetic resonance (NMR) is a spectroscopic technique that allows researchers to study the structure and dynamics of molecules at atomic resolution. It has developed up to the point where it plays a vital role in analytical- and structural research and even contributes to medical screening, material sciences and micro-imaging (microMRI). The wide range of applications showcases the flexibility of NMR and, due to its non-invasive nature, its large potential in biology. Recent advances in sample preparation now even make it possible to study proteins at atomic details in situ or even directly in cells (cellular NMR)[30,109,143].

In short, NMR makes use of magnetic property of nuclear spins, which start to precess around the axis of an external magnetic field. The precession of a spin is also known as resonance frequency, which is unique for each individual nuclei inside a sample. These resonance frequencies are extremely sensitive for local changes in the chemical environment and can therefore be used to study the structure and structural changes of biological samples e.g. of a protein. Moreover, the dynamics can be studied from picosecond up to second timescale. In contrast to solution NMR, where fast tumbling smaller proteins/complexes can be studied, solid state NMR spectroscopy is less limited by molecular size enabling studies of biological samples in their native (often dense) physiological environment.

Fundamentals in ssNMR

Biological NMR relies on NMR-active elements which are often ^{1}H, ^{2}H, ^{13}C, ^{15}N, 19F and ^{31}P.

A common feature for these nuclei relies on the property called spin (I), which describes the energy states when placed in a magnetic field. Biologically relevant nuclei often display an I-value of ½. NMR makes use of the spin state transition energy which is given by eq.1:

$$\Delta E = \hbar \gamma B_0$$

where ΔE is the energy difference between the two relevant spin states, $\hbar$ is reduced Plank's constant, γ is the gyromagnetic ratio and B_0 is the magnitude of the external magnetic field. When nuclei are placed inside a magnetic field, the spins will orient themselves parallel or antiparallel with the direction of the magnetic field. The distribution of these energy states (N_h and N_l) can be described using the Boltzmann Equation (eq.2; Fig 3.a):

$$\frac{N_h}{N_l} = e^{-\gamma \hbar B_0 / T}$$

where B_0 is the magnitude of the external magnetic field, γ is the gyromagnetic ratio, $\hbar$ is reduced Plank's constant and T denotes the temperature. The gyromagnetic ratio (γ_n) is an intrinsic property of a nucleus dependent on its relationship between its magnetic moment and angular momentum following eq.3:

$$\gamma_n = \frac{e}{2m_p} g_n$$

where m_p is the proton rest mass and g_n is the g-factor of that specific nucleus.

Two nuclei can interact with one another through various types of interactions. These interactions are described by a Hamiltonian that defines the resonating energy of a nucleus, given by the spin interactions with the external B_0 giving rise to a chemical shift. Interactions between two spins (I and S) are commonly described by the scalar (also known as J) and dipolar coupling. The corresponding Hamiltonians are given in table 1.

Spin interaction type	Hamiltonian
Chemical shift	$\hat{\mathrm{H}}_{cs} = \gamma_I \hat{\mathrm{I}} \cdot \sigma \cdot B_0$
J-Coupling	$\hat{\mathrm{H}}_j = J \cdot \hat{\mathrm{I}} \cdot \hat{\mathrm{S}}$
Homonuclear D-D coupling	$\hat{\mathrm{H}}_{DD} = -\frac{\mu_0 \gamma_I \gamma_S \hbar}{4\pi \cdot r_{I,S}^3} \left(\frac{1}{2} 3cos^2\theta - 1\right) \left(3\hat{\mathrm{I}}_z \cdot \hat{\mathrm{S}}_z - \hat{\mathrm{I}} \cdot \hat{\mathrm{S}}\right)$
Heteronuclear D-D coupling	$\hat{\mathrm{H}}_{DD} = -\frac{\mu_0 \gamma_I \gamma_S \hbar}{4\pi \cdot r_{I,S}^3} (3cos^2\theta - 1)\left(2\hat{\mathrm{I}}_z \cdot \hat{\mathrm{S}}_z\right)$

Table 1: Overview of Hamiltonians which describe spin interactions of spin I and S.
Definitions: $\hat{H}$ Hamiltonian, γ gyromagnetic ratio, σ chemical shift tensor, B_0 magnitude of the external magnetic field, J J-coupling constant, μ vacuum permeability, $\hbar$ reduced Plank's constant, r distance between two nuclei and θ is the angle between the two nuclei and the external magnetic field.

The external magnetic field experienced by a NMR-active nucleus depends on its position relative to the direction of the static magnetic field. Moreover, local interactions with surrounding electrons will also influence the magnetic field experienced by the nucleus, termed "chemical shielding". Together, they indirectly induce a shift of the signal in a NMR spectrum. Molecules are normally randomly distributed inside the magnetic field, which results in each of them resonating to a unique frequency leading to an ensemble of chemical shifts and giving rise in a static case to a so-called "Powder spectrum", which is defined by the chemical shift anisotropy (CSA) (Fig 3.b,c).

The isotropic J-coupling or scalar coupling, is an interaction between two nuclei which are connected by covalent chemical bonds. The interaction originates from the hyperfine couplings between the nuclei and electrons. The electrons assist the coupling between the two nuclei and is therefore called indirect. As the couplings depend on the through-bound interactions, it is exclusively intramolecular and often gives accurate information on the connectivity of chemical bonds and provides additional information e.g. bond lengths or angles.

In contrast to J-couplings, nuclei can also interact through space via dipolar couplings (Fig 3.d). Each individual spin has a small magnetic field, which interacts with a second nuclear spin and the dipolar coupling interaction is therefore mutual between the two spins. Additionally, dipolar interactions can be both intra- as intermolecular. These couplings are not observed in solution NMR as the rapid tumbling of the studied molecules averages these interactions to zero. In ssNMR, dipole-dipole interactions generate high amounts of line broadening, as the molecules naturally hardly tumble and the NMR spectrum is a superposition of all orientations of the dipolar vector in the static magnetic field (Figure 3d). Dipolar couplings interact through space and are sensitive towards the distance of

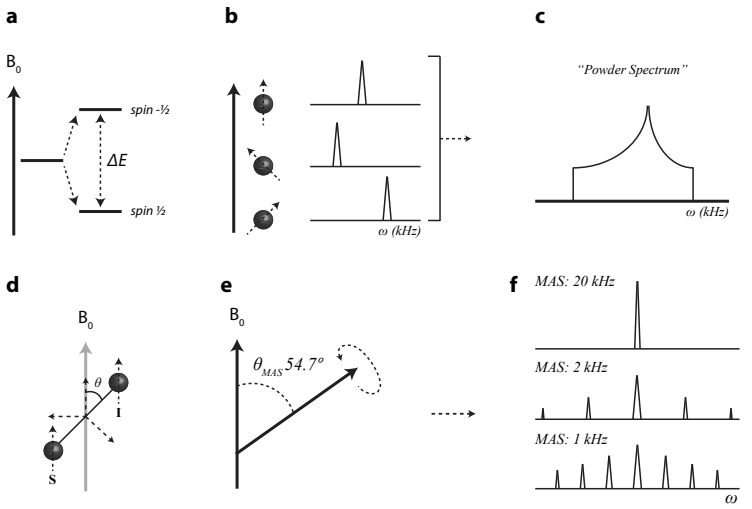


Figure 3 | Overview of ssNMR fundamentals. a, representation of Zeeman energy levels for spin ½ and spin -½ when an external B0 magnetic field is applied. **b**, the chemical shift is dependent of the orientation of the molecule in respect to the magnetic field which, **c**, leads to accumulation of an ensemble of peaks resulting in a powder spectrum. **d,** the strength of dipole-dipole interaction depends on the distance of the two nuclei in space and their orientation in respect to the B0 field. **e**, spinning the sample at the frequency of the dipolar couplings around the magic angle will increase, **f**, the spectral resolution.

interacting nuclei. Therefore, they frequently are used to obtain distance information and restraints in ssNMR.

As discussed in the section above, the combination of amorphous samples and strong dipolar couplings in solids, the general spectra is of low quality, i.e., limited spectral resolution. To improve the resolution of the spectra, the sample is artificially spun around the magic angle (Magic Angle Spinning: MAS). The dipole-dipole Hamiltonian expresses a term which includes the angle of the internuclear vector and the external magnetic field. The dipole-dipole coupling can be solved for zero by adjusting the angle around which the sample is spun (Fig 3.e,f):

$$3cos^2\theta - 1 = 0$$

$$\theta_{magic} = \arctan\sqrt{2} \cong 54.74^\circ$$

The MAS frequency at which samples are spun depends on the size of the dipolar coupling desired to be removed. The spinning frequency needs to be in the order of magnitude of the dipolar interaction to be averaged out which amounts to about 5 kHz for ^{13}C-^{13}C dipolar couplings (with an interuclear distance of 1.5 Angstrom) or 30 kHz for one-bond ^{13}C-^{1}H couplings. To miminize broadening effects in the ssNMR spectrum, detection is usually performed on ^{13}C or on ^{15}N, during which the ^{1}H couplings (~120 kHz) are decoupled during detection. Recent developments in ssNMR probes and advanced labelling schemes that reduce the amount of ^{1}H (proton) spins in the sample have opened up new avenues of detection methods. Today, such probes allow to use ultra-fast MAS frequencies ranging for 60 kHz for a 1.3 mm MAS rotor up to 100-111 kHz for 0.7 mm rotors. As mentioned above, the ^{1}H-^{1}H couplings can also be lowered by reducing the protonation level and breaking the ^{1}H-^{1}H network using biochemical approached.

One of the major drawbacks of NMR is its low sensitivity, as the population difference of N_h and N_l discussed in eq. 2 under normal conditions is small. As described in eq.1, the theoretical maximum of a signal is proportional to the distribution of energy states can be affected given by the Boltzmann equation. This implies that the signal scales with increasing magnetic fields and higher gyromagnetic ratios. In biological ssNMR, ^{13}C and ^{15}N are the most commonly used nuclei to detect. However, these nuclei have respectively a 4-times and 10-times smaller gyromagnetic ratio as compared to ^{1}H. One way of utilizing this higher gyromagnetic ratio is by using cross-polarization (CP) experiments, where one can increase the sensitivity of e.g. ^{13}C and ^{15}N. The principle of CP relies on the dipolar interactions that allow to transfer the magnetization on ^{1}H to another nucleus (such as ^{13}C or ^{15}N) provided that the Hartmann-Hahn CP condition, in which nuclei are irradiated at the same time (eq.5)[144,145] is fulfilled:

$$\omega_X = \omega_{1_H} \pm n\omega_{MAS}\,, with\ n = (1,2)$$

Another indirect advantage is that by using ^{1}H during CP, the T_1 time of the experiment shortens, allowing for an increased number of scans during the same measurement period.

A second way of utilizing the high gyromagnetic ratio of ^{1}H is by directly detecting these nuclei[146,147]. With the development of ultra-fast MAS probes and rotors, researchers can now spin their samples at frequencies close to that of ^{1}H-^{1}H couplings[148]. Resolution can be further supported by implementing deuteration strategies to further disrupt proton networks[149].

Finally, if the use of proton-detection is not sensitive enough, one can resort to using the gyromagnetic ratio of electrons as it is 658-times higher than the gyromagnetic ratio of protons. This specialized technique called Dynamic Nuclear Polarization (DNP) uses microwaves generated by a gyrotron to irradiate the sample[150,151]. The microwaves polarize the free electron pairs on tailored bi-radicals, e.g. TOTAPOL or AMUPol[152], which are present in the buffer of the sample. When the conditions for so called cross-effect DNP are met[153], the magnetization can be transferred from the electrons in the bi-radicals to the protons of the protein of interest (eq.6):

$$\omega_{nucleus} = \omega_{e1} \pm \omega_{e2}$$

In ideal situations, NMR signal enhancements can be realized of more than 100-fold, reducing NMR experiment times by 100^2 =10.000 times. The interaction of the electron with other nuclei is also anisotropic, which means that the efficiency of magnetization transfer is dependent on the orientation of the molecules and will therefore be modulated by MAS. Moreover, the external magnetic field strength also influences the maximal enhancement that can be obtained, with higher magnetic fields reducing the enhancement for biradicals such AMUPol. Even though DNP allows the study of low molecule concentrations, it has one major limitation currently refers to the fact that experiments typically are performed at 100K. Freezing samples can give rise to structural heterogeneity leading to enhanced line broadening in the spectra compared to ssNMR at ambient temperatures.

ssNMR Experiments

As the section above describes, the combination of rigid/folded protein domains and motion are crucially important in the structure-function relationship of proteins. ssNMR is a unique technique which allows researchers to study these principles under physiological conditions via motion-sensitive (J-based) or rigidity-sensitive (dipolar-based) experiment. The section below will discuss some of the most frequently used ssNMR experiments in this thesis.

In conventional ssNMR, nuclei with lower gyromagnetic ratio are more frequently used for detection. Several scalar-based ssNMR pulse sequences have been developed to study homonuclear (^{13}C-^{13}C; INEPT TOBSY[154]) or heteronuclear (^{1}H-^{13}C; INEPT HETCOR (Fig 4.a)) J-couplings. In short, protons are irradiated in the protein after which the magnetization is transferred through scalar couplings to connecting ^{13}C, after which, in the case of the INEPT TOBSY, relayed to another chemically bonded ^{13}C. In both cases, detection is performed on ^{13}C while simultaneously the protons are decoupled. Alternatively, one could use HSQC-type experiments as used in solution-state NMR, where protons are directly detected. Here, polarization is transferred through-bond to nearby X-nuclei (^{13}C or ^{15}N) and back again to proton for its detection.

The advantage of using the pulse sequences is that there is also a dipolar-based experiment which probes the same nuclei. For example, ^{13}C-^{13}C Proton-Driven Spin Diffusion (PDSD[155]) follows the same principle as an INEPT TOBSY experiment, but the transfer of magnetization from proton to ^{13}C and from ^{13}C to ^{13}C is realized via dipolar couplings (Fig 4.b). The INEPT TOBSY and PDSD probe unique protein domains as higher degrees of motion, e.g. unfolded protein segments, are easily observed by an INEPT TOBSY while folded protein domains (that exhibit higher rigidity) are readily detected in a dipolar-based PDSD. As the dipolar couplings interact through space, PDSD also contains distance

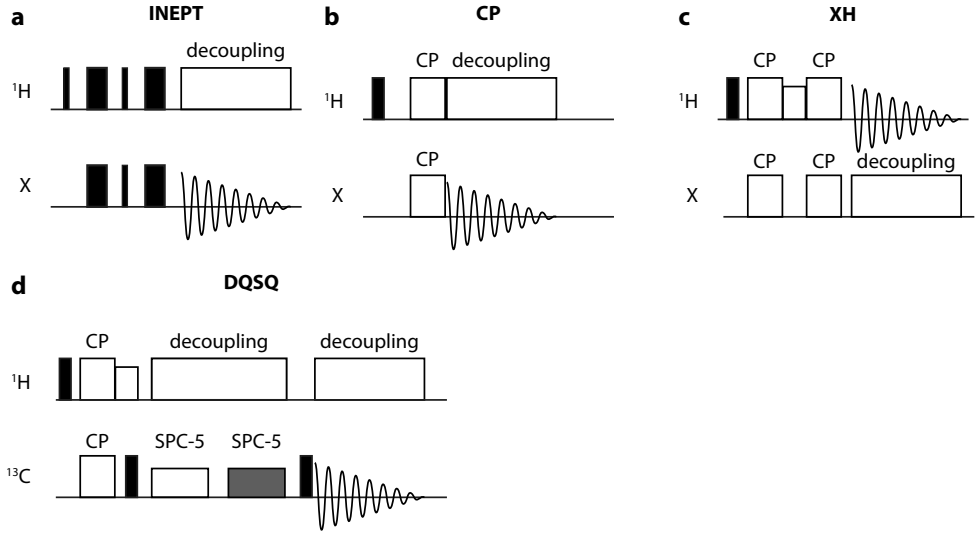


Figure 4 | Pulse sequences of frequently used ssNMR experiments in this thesis. Overview of ssNMR pulse sequences which are scalar-based, **a**, or dipolar-based (**b-d**). Solid boxes indicate 90° pulses. **c**, A water suppression sequence incorporated between the two CP steps during proton-detection suppress the water signal.

information where cross-peaks are observed up to 5Å. Also dipolar-based XH experiments are developed, which follow the same principles as the scalar counterparts, albeit the magnetization is transferred through space (Fig 4.c).

Finally, two chapters in this thesis make extensive use of Double Quantum Single Quantum (DQSQ) experiments[156] (Fig 4.d). These experiments are dipolar-based and probe ^{13}C-^{13}C correlations. DQSQ pulse sequences make use of double quantum recoupling scheme such as the SPC5 sequence[156]. Such sequence reintroduce the dipolar coupling that, as discussed above, are averaged out MAS conditions, among ^{13}C spins by synchronizing the pulse sequence to the MAS-induced rotation of the sample. In detail, radio-frequency (r.f.) pulses are applied with an r.f. amplitude that usually is am multiple of the MAS frequency. During SPC5 recoupling, the efficiency of the recoupling is sensitive to the motion of the molecule: Dipolar recoupling occurs more efficiently when molecules are more rigid[157] while local motion leads to reduction of the NMR signal. Therefore, these spectra can be used to perform comparative studies where changes in protein dynamics lead to changes in the signal intensity[30].

SCOPE OF THE THESIS

In this thesis we use established protocols for cellular ssNMR in order to study the role of protein motion in growth factor receptor activation and functioning. In addition, novel approaches are discussed in the context of studying protein phase separation and develop new methods are introduced that push the boundaries of native ssNMR. **Chapter 2** provides a literature overview of recent developments in native and cellular NMR for bacterial- and mammalian systems and highlights the strengths and limitations of in-cell NMR. In **chapter 3**,

we study the role of conformational dynamics in the molecular activation and functioning of full-length EGFR in its native plasma membrane. This work is further supported by **chapter 4,** where a family member of EGFR, Her2, is also studied in situ. The latter two chapters illustrate that the molecular activation mechanism differs from the case of using *in vitro* models, even if the in situ studies are performed in isolated membranes. Therefore, we set out to develop a new method in **chapter 5**, where we grow 3-dimensional cell preparations for ssNMR studies which could be used as novel system for structural studies in in vitro-grown tumor models. Finally, in **chapter 6** we describe a unique approach where we use a combination of solution- and solid-state NMR to elucidate the macromolecular structure of proteins after phase separation. The thesis ends with a concise summary of our findings.

REFERENCES

1. Zhu, S. et al. Mechanism of NMDA Receptor Inhibition and Activation. Cell (2016). doi:10.1016/j.cell.2016.03.028
2. Tajima, N. et al. Activation of NMDA receptors and the mechanism of inhibition by ifenprodil. Nature (2016). doi:10.1038/nature17679
3. Karakas, E. & Furukawa, H. Crystal structure of a heterotetrameric NMDA receptor ion channel. Science (80-.). (2014). doi:10.1126/science.1251915
4. Nygaard, R. et al. The dynamic process of β2-adrenergic receptor activation. Cell (2013). doi:10.1016/j.cell.2013.01.008
5. Staus, D. P. et al. Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. Nature (2016). doi:10.1038/nature18636
6. Stuttfeld, E. & Ballmer-Hofer, K. Structure and function of VEGF receptors. IUBMB Life (2009). doi:10.1002/iub.234
7. Leppänen, V. M. et al. Structural and mechanistic insights into VEGF receptor 3 ligand binding and activation. Proc. Natl. Acad. Sci. U. S. A. (2013). doi:10.1073/pnas.1301415110
8. Simons, M., Gordon, E. & Claesson-Welsh, L. Mechanisms and regulation of endothelial VEGF receptor signalling. Nature Reviews Molecular Cell Biology (2016). doi:10.1038/nrm.2016.87
9. Bose, R. A neu view of invasive lobular breast cancer. Clinical Cancer Research (2013). doi:10.1158/1078-0432.CCR-13-1031
10. Ross, J. S. et al. Relapsed classic E-cadherin (CDH1)-mutated invasive lobular breast cancer shows a high frequency of HER2 (ERBB2) gene mutations. Clin. Cancer Res. (2013). doi:10.1158/1078-0432.CCR-13-0295
11. Salomon, D. S., Brandt, R., Ciardiello, F. & Normanno, N. Epidermal growth factor-related peptides and their receptors in human malignancies. Critical Reviews in Oncology and Hematology (1995). doi:10.1016/1040-8428(94)00144-I
12. Herbst, R. S. & Shin, D. M. Monoclonal antibodies to target epidermal growth factor receptor-positive tumors a new paradigm for cancer therapy. Cancer (2002). doi:10.1002/cncr.10372
13. Messa, C., Russo, F., Caruso, M. G. & Di Leo, A. EGF, TGF-□, and EGF-R in human colorectal adenocarcinoma. Acta Oncol. (Madr). (1998). doi:10.1080/028418698429595
14. Ogiso, H. et al. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell (2002). doi:S0092867402009637 [pii]
15. Ferguson, K. M. et al. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol. Cell (2003). doi:10.1016/S1097-2765(03)00047-9
16. Ferguson, K. M. Structure-Based View of Epidermal Growth Factor Receptor Regulation. Annu. Rev. Biophys. (2008). doi:10.1146/annurev.biophys.37.032807.125829
17. Yu, X., Sharma, K. D., Takahashi, T., Iwamoto, R. & Mekada, E. Ligand-independent dimer formation of epidermal growth factor receptor (EGFR) is a step separable from ligand-induced EGFR signaling. Mol. Biol. Cell (2002). doi:10.1091/mbc.01-08-0411
18. Low-Nam, S. T. et al. ErbB1 dimerization is promoted by domain co-confinement and stabilized by ligand binding. Nat. Struct. Mol. Biol. (2011). doi:10.1038/nsmb.2135
19. Freed, D. M., Alvarado, D. & Lemmon, M. A. Ligand regulation of a constitutively dimeric EGF receptor. Nat. Commun. (2015). doi:10.1038/ncomms8380
20. Chung, I. et al. Spatial control of EGF receptor activation by reversible dimerization on living cells. Nature (2010). doi:10.1038/nature08827
21. Saffarian, S., Li, Y., Elson, E. L. & Pikey, L. J. Oligomerization of the EGF receptor investigated by live cell fluorescence intensity distribution analysis. Biophys. J. (2007). doi:10.1529/biophysj.107.105494
22. Bader, A. N., Hofman, E. G., Voortman, J., Van Bergen En Henegouwen, P. M. P. & Gerritsen, H. C. Homo-FRET imaging enables quantification of protein cluster sizes with subcellular resolution. Biophys. J. (2009). doi:10.1016/j.bpj.2009.07.059
23. Zanetti-Domingues, L. C. et al. The architecture of EGFR's basal complexes reveals autoinhibition mechanisms in dimers and oligomers. Nat. Commun. (2018). doi:10.1038/s41467-018-06632-0

24. Moriki, T., Maruyama, H. & Maruyama, I. N. Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain. J. Mol. Biol. (2001). doi:10.1006/jmbi.2001.4923
25. Jura, N. et al. Mechanism for Activation of the EGF Receptor Catalytic Domain by the Juxtamembrane Segment. Cell (2009). doi:10.1016/j.cell.2009.04.025
26. Arkhipov, A. et al. Architecture and membrane interactions of the EGF receptor. Cell (2013). doi:10.1016/j.cell.2012.12.030
27. Alvarado, D., Klein, D. E. & Lemmon, M. A. Structural Basis for negative cooperativity in growth factor binding to an EGF receptor. Cell (2010). doi:10.1016/j.cell.2010.07.015
28. Lu, C., Mi, L. Z., Schürpf, T., Walz, T. & Springer, T. A. Mechanisms for kinase-mediated dimerization of the epidermal growth factor receptor. J. Biol. Chem. (2012). doi:10.1074/jbc.M112.414391
29. Bessman, N. J., Bagchi, A., Ferguson, K. M. & Lemmon, M. A. Complex Relationship between Ligand Binding and Dimerization in the Epidermal Growth Factor Receptor. Cell Rep. (2014). doi:10.1016/j.celrep.2014.10.010
30. Kaplan, M. et al. EGFR Dynamics Change during Activation in Native Membranes as Revealed by NMR. Cell (2016). doi:10.1016/j.cell.2016.10.038
31. Citri, A. & Yarden, Y. EGF-ERBB signalling: Towards the systems level. Nature Reviews Molecular Cell Biology (2006). doi:10.1038/nrm1962
32. Schneider, M. R. & Wolf, E. The epidermal growth factor receptor ligands at a glance. Journal of Cellular Physiology (2009). doi:10.1002/jcp.21635
33. Kovacs, E., Zorn, J. A., Huang, Y., Barros, T. & Kuriyan, J. A Structural Perspective on the Regulation of the Epidermal Growth Factor Receptor. Annu. Rev. Biochem. (2015). doi:10.1146/annurev-biochem-060614-034402
34. Freed, D. M. et al. EGFR Ligands Differentially Stabilize Receptor Dimers to Specify Signaling Kinetics. Cell (2017). doi:10.1016/j.cell.2017.09.017
35. Cho, H. S. et al. Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature (2003). doi:10.1038/nature01392
36. Garrett, T. P. J. et al. The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors. Mol. Cell (2003). doi:10.1016/S1097-2765(03)00048-0
37. Alvarado, D., Klein, D. E. & Lemmon, M. A. ErbB2 resembles an autoinhibited invertebrate epidermal growth factor receptor. Nature (2009). doi:10.1038/nature08297
38. Burgess, A. W. et al. An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Molecular Cell (2003). doi:10.1016/S1097-2765(03)00350-2
39. Heukers, R. et al. Endocytosis of EGFR requires its kinase activity and N-terminal transmembrane dimerization motif. J. Cell Sci. (2013). doi:10.1242/jcs.128611
40. Mineev, K. S., Panova, S. V., Bocharova, O. V., Bocharov, E. V. & Arseniev, A. S. The Membrane Mimetic Affects the Spatial Structure and Mobility of EGFR Transmembrane and Juxtamembrane Domains. Biochemistry (2015). doi:10.1021/acs.biochem.5b00851
41. Bocharov, E. V. et al. Alternative packing of EGFR transmembrane domain suggests that protein-lipid interactions underlie signal conduction across membrane. Biochim. Biophys. Acta - Biomembr. (2016). doi:10.1016/j.bbamem.2016.02.023
42. McLaughlin, S., Smith, S. O., Hayman, M. J. & Murray, D. An electrostatic engine model for autoinhibition and activation of the epidermal growth factor receptor (EGFR/ErbB) family. J. Gen. Physiol. (2005). doi:10.1085/jgp.200509274
43. Matsushita, C. et al. Transmembrane helix orientation influences membrane binding of the intracellular juxtamembrane domain in Neu receptor peptides. Proc. Natl. Acad. Sci. U. S. A. (2013). doi:10.1073/pnas.1215207110
44. Baker, L. A. & Baldus, M. Characterization of membrane protein function by solid-state NMR spectroscopy. Current Opinion in Structural Biology (2014). doi:10.1016/j.sbi.2014.03.009
45. Smith, S. O. et al. Transmembrane interactions in the activation of the Neu receptor tyrosine kinase. Biochemistry (2002). doi:10.1021/bi012117l
46. Doerner, A., Scheck, R. & Schepartz, A. Growth Factor Identity Is Encoded by Discrete Coiled-Coil Rotamers in the EGFR Juxtamembrane Region. Chem. Biol. (2015). doi:10.1016/j.chembiol.2015.05.008

47. Sinclair, J. K. L., Walker, A. S., Doerner, A. E. & Schepartz, A. Mechanism of Allosteric Coupling into and through the Plasma Membrane by EGFR. Cell Chem. Biol. (2018). doi:10.1016/j.chembiol.2018.04.005
48. Yun, C. H. et al. Structures of Lung Cancer-Derived EGFR Mutants and Inhibitor Complexes: Mechanism of Activation and Insights into Differential Inhibitor Sensitivity. Cancer Cell (2007). doi:10.1016/j.ccr.2006.12.017
49. Wood, E. R. et al. A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): Relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. Cancer Res. (2004). doi:10.1158/0008-5472.CAN-04-1168
50. Stamos, J., Sliwkowski, M. X. & Eigenbrot, C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. J. Biol. Chem. (2002). doi:10.1074/jbc.M207135200
51. Jura, N. et al. Catalytic Control in the EGF Receptor and Its Connection to General Kinase Regulatory Mechanisms. Molecular Cell (2011). doi:10.1016/j.molcel.2011.03.004
52. Chen, H. et al. A Molecular Brake in the Kinase Hinge Region Regulates the Activity of Receptor Tyrosine Kinases. Mol. Cell (2007). doi:10.1016/j.molcel.2007.06.028
53. Shigematsu, H. & Gazdar, A. F. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. International Journal of Cancer (2006). doi:10.1002/ijc.21496
54. Sutto, L. & Gervasio, F. L. Effects of oncogenic mutations on the conformational free-energy landscape of EGFR kinase. Proc. Natl. Acad. Sci. U. S. A. (2013). doi:10.1073/pnas.1221953110
55. Gajiwala, K. S. et al. Insights into the aberrant activity of mutant EGFR kinase domain and drug recognition. Structure (2013). doi:10.1016/j.str.2012.11.014
56. Mustafa, M., Mirza, A. & Kannan, N. Conformational regulation of the EGFR kinase core by the juxtamembrane and C-terminal tail: A molecular dynamics study. Proteins Struct. Funct. Bioinforma. (2011). doi:10.1002/prot.22862
57. De Heus, C., Kagie, N., Heukers, R., Van Bergen En Henegouwen, P. M. P. & Gerritsen, H. C. Analysis of EGF Receptor Oligomerization by Homo-FRET. in Methods in Cell Biology (2013). doi:10.1016/B978-0-12-408143-7.00016-5
58. Nagy, P. et al. Activation-dependent clustering of the erbB2 receptor tyrosine kinase detected by scanning near-field optical microscopy. J. Cell Sci. (1999).
59. Tobin, S. J. et al. Single molecule localization microscopy coupled with touch preparation for the quantification of trastuzumab-bound HER2. Sci. Rep. (2018). doi:10.1038/s41598-018-33225-0
60. Huang, Y. et al. Molecular basis for multimerization in the activation of the epidermal growth factor receptor. Elife (2016). doi:10.7554/eLife.14107
61. Needham, S. R. et al. EGFR oligomerization organizes kinase-active dimers into competent signalling platforms. Nat. Commun. (2016). doi:10.1038/ncomms13307
62. Sorkin, A. & Goh, L. K. Endocytosis and intracellular trafficking of ErbBs. Experimental Cell Research (2009). doi:10.1016/j.yexcr.2008.07.029
63. Carpentier, J. L. et al. Co-localization of 125I-epidermal growth factor and ferritin-low density lipoprotein in coated pits: A quantitative electron microscopic study in normal mutant human fibroblasts. J. Cell Biol. (1982). doi:10.1083/jcb.95.1.73
64. Wiley, H. S. Anomalous binding of epidermal growth factor to A431 cells is due to the effect of high receptor densities and a saturable endocytic system. J. Cell Biol. (1988). doi:10.1083/jcb.107.2.801
65. Sigismund, S. et al. Clathrin-independent endocytosis of ubiquitinated cargos. Proc. Natl. Acad. Sci. U. S. A. (2005). doi:10.1073/pnas.0409817102
66. Sigismund, S. et al. Clathrin-Mediated Internalization Is Essential for Sustained EGFR Signaling but Dispensable for Degradation. Dev. Cell (2008). doi:10.1016/j.devcel.2008.06.012
67. Lu, C. et al. Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal growth factor receptor. Mol. Cell. Biol. (2010). doi:10.1128/MCB.00742-10
68. Heukers, R. et al. Endocytosis of EGFR requires its kinase activity and N-terminal transmembrane dimerization motif. J. Cell Sci. (2013). doi:10.1242/jcs.128611
69. Ariotti, N. et al. Epidermal Growth Factor Receptor Activation Remodels the Plasma Membrane Lipid

Environment To Induce Nanocluster Formation. Mol. Cell. Biol. (2010). doi:10.1128/mcb.01615-09

70. Jones, D. H., Rigby, A. C., Barber, K. R. & Grant, . W. M. Oligomerization of the EGF receptor transmembrane domain: A 2H NMR study in lipid bilayers. Biochemistry (1997). doi:10.1021/bi970547z
71. Popov-Čeleketic, D. & van Bergen en Henegouwen, P. M. P. Membrane domain formation - A key factor for targeted intracellular drug delivery. Frontiers in Physiology (2014). doi:10.3389/fphys.2014.00462
72. Henne, W. M. et al. FCHo proteins are nucleators of Clathrin-Mediated endocytosis. Science (80-.). (2010). doi:10.1126/science.1188462
73. McMahon, H. T. & Boucrot, E. Molecular mechanism and physiological functions of clathrin-mediated endocytosis. Nature Reviews Molecular Cell Biology (2011). doi:10.1038/nrm3151
74. Taylor, M. J., Perrais, D. & Merrifield, C. J. A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis. PLoS Biol. (2011). doi:10.1371/journal.pbio.1000604
75. Orellana, L. et al. Oncogenic mutations at the EGFR ectodomain structurally converge to remove a steric hindrance on a kinase-coupled cryptic epitope. Proc. Natl. Acad. Sci. U. S. A. (2019). doi:10.1073/pnas.1821442116
76. Lin, Y., Protter, D. S. W., Rosen, M. K. & Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol. Cell (2015). doi:10.1016/j.molcel.2015.08.018
77. Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: Modular design for efficient function. Nature Reviews Molecular Cell Biology (2007). doi:10.1038/nrm2178
78. Boke, E. et al. Amyloid-like Self-Assembly of a Cellular Compartment. Cell (2016). doi:10.1016/j.cell.2016.06.051
79. Decker, C. J., Teixeira, D. & Parker, R. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. J. Cell Biol. (2007). doi:10.1083/jcb.200704147
80. Kwon, I. et al. Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. Science (80-.). (2014). doi:10.1126/science.1254917
81. Conicella, A. E., Zerze, G. H., Mittal, J. & Fawzi, N. L. ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. Structure (2016). doi:10.1016/j.str.2016.07.007
82. Woodruff, J. B. et al. The Centrosome Is a Selective Condensate that Nucleates Microtubules by Concentrating Tubulin. Cell (2017). doi:10.1016/j.cell.2017.05.028
83. Feric, M. et al. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell (2016). doi:10.1016/j.cell.2016.04.047
84. Su, X. et al. Phase separation of signaling molecules promotes T cell receptor signal transduction. Science (80-.). (2016). doi:10.1126/science.aad9964
85. Sheu-Gruttadauria, J. & MacRae, I. J. Phase Transitions in the Assembly and Function of Human miRISC. Cell (2018). doi:10.1016/j.cell.2018.02.051
86. Fromm, S. A. et al. The structural basis of Edc3-and Scd6-mediated activation of the Dcp1:Dcp2 mRNA decapping complex. EMBO J. (2012). doi:10.1038/emboj.2011.408
87. Fromm, S. A. et al. InVitro reconstitution of a cellular phase-transition process that involves the mRNA decapping machinery. Angew. Chemie - Int. Ed. (2014). doi:10.1002/anie.201402885
88. Schutz, S., Noldeke, E. R. & Sprangers, R. A synergistic network of interactions promotes the formation of in vitro processing bodies and protects mRNA against decapping. Nucleic Acids Res. (2017). doi:10.1093/nar/gkx353
89. Franzmann, T. M. et al. Phase separation of a yeast prion protein promotes cellular fitness. Science (80-.). (2018). doi:10.1126/science.aao5654
90. Kroschwald, S. et al. Different Material States of Pub1 Condensates Define Distinct Modes of Stress Adaptation and Recovery. Cell Rep. (2018). doi:10.1016/j.celrep.2018.05.041
91. Riback, J. A. et al. Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. Cell (2017). doi:10.1016/j.cell.2017.02.027
92. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. Cell (2019). doi:10.1016/j.cell.2018.12.035
93. Schmidt, H. B. & Görlich, D. Transport Selectivity of

Nuclear Pores, Phase Separation, and Membraneless Organelles. Trends in Biochemical Sciences (2016). doi:10.1016/j.tibs.2015.11.001

94. Brangwynne, C. P. Soft active aggregates: Mechanics, dynamics and self-assembly of liquid-like intracellular protein bodies. Soft Matter (2011). doi:10.1039/c0sm00981d
95. Brangwynne, C. P. Phase transitions and size scaling of membrane-less organelles. J. Cell Biol. (2013). doi:10.1083/jcb.201308087
96. Patel, A. et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell (2015). doi:10.1016/j.cell.2015.07.047
97. Boeynaems, S. et al. Protein Phase Separation: A New Phase in Cell Biology. Trends in Cell Biology (2018). doi:10.1016/j.tcb.2018.02.004
98. Wu, H. & Fuxreiter, M. The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. Cell (2016). doi:10.1016/j.cell.2016.05.004
99. March, Z. M., King, O. D. & Shorter, J. Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. Brain Research (2016). doi:10.1016/j.brainres.2016.02.037
100. Alberti, S., Halfmann, R., King, O., Kapila, A. & Lindquist, S. A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins. Cell (2009). doi:10.1016/j.cell.2009.02.044
101. Hughes, M. P. et al. Atomic structures of low-complexity protein segments reveal kinked b sheets that assemble networks. Science (80-.). (2018). doi:10.1126/science.aan6398
102. Sawaya, M. R. et al. Atomic structures of amyloid cross-β spines reveal varied steric zippers. Nature (2007). doi:10.1038/nature05695
103. Murray, D. T. et al. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. Cell (2017). doi:10.1016/j.cell.2017.08.048
104. Murthy, A. C. et al. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. Nat. Struct. Mol. Biol. (2019). doi:10.1038/s41594-019-0250-x
105. Ackermann, B. E. & Debelouchina, G. T. Heterochromatin Protein HP1α Gelation Dynamics Revealed by Solid-State NMR Spectroscopy. Angew. Chemie - Int. Ed. (2019). doi:10.1002/anie.201901141
106. Ader, C. et al. Amyloid-like interactions within nucleoporin FG hydrogels. Proc. Natl. Acad. Sci. U. S. A. (2010). doi:10.1073/pnas.0910163107
107. Baker, L. A. et al. Combined 1H-Detected Solid-State NMR Spectroscopy and Electron Cryotomography to Study Membrane Proteins across Resolutions in Native Environments. Structure (2018). doi:10.1016/j.str.2017.11.011
108. Visscher, K. M. et al. Supramolecular Organization and Functional Implications of K + Channel Clusters in Membranes. Angew. Chemie - Int. Ed. (2017). doi:10.1002/anie.201705723
109. Medeiros-Silva, J. et al. High-resolution NMR studies of antibiotics in cellular membranes. Nat. Commun. (2018). doi:10.1038/s41467-018-06314-x
110. Özcan, F., Klein, P., Lemmon, M. A., Lax, I. & Schlessinger, J. On the nature of low- and high-affinity EGF receptors on living cells. Proc. Natl. Acad. Sci. U. S. A. (2006). doi:10.1073/pnas.0601469103
111. Miljan, E. A. et al. Interaction of the extracellular domain of the epidermal growth factor receptor with gangliosides. J. Biol. Chem. (2002). doi:10.1074/jbc.M111669200
112. Pike, L. J. & Casey, L. Cholesterol levels modulate EGF receptor-mediated signaling by altering receptor function and trafficking. Biochemistry (2002). doi:10.1021/bi025943i
113. Orr, G. et al. Cholesterol dictates the freedom of EGF receptors and HER2 in the plane of the membrane. Biophys. J. (2005). doi:10.1529/biophysj.104.056192
114. Guan, F., Handa, K. & Hakomori, S. I. Regulation of epidermal growth factor receptor through interaction of ganglioside GM3 with GlcNAc of N-Linked glycan of the receptor: Demonstration in ldlD cells. Neurochem. Res. (2011). doi:10.1007/s11064-010-0379-9
115. Hanai, N., Nores, G. A., MacLeod, C., Torres-Mendez, C. R. & Hakomori, S. Ganglioside-mediated modulation of cell growth. Specific effects of GM3 and lyso-GM3 in tyrosine phosphorylation of the epidermal growth factor receptor. J. Biol. Chem. (1988).
116. Zhou, Q., Hakomori, S. I., Kitamura, K. & Igarashi,

Y. G(M3) directly inhibits tyrosine phosphorylation and de-N-acetyl-G(M3) directly enhances serine phosphorylation of epidermal growth factor receptor, independently of receptor-receptor interaction. J. Biol. Chem. (1994).

117. Wang, X. Q., Sun, P. & Paller, A. S. Ganglioside induces caveolin-1 redistribution and interaction with the epidermal growth factor receptor. J. Biol. Chem. (2002). doi:10.1074/jbc.M208257200

118. Coskun, Ü., Grzybek, M., Drechsel, D. & Simons, K. Regulation of human EGF receptor by lipids. Proc. Natl. Acad. Sci. U. S. A. (2011). doi:10.1073/pnas.1105666108

119. Hofman, E. G. et al. EGF induces coalescence of different lipid rafts. J. Cell Sci. (2008). doi:10.1242/jcs.028753

120. Hakomori, S. I. & Zhang, Y. Glycosphingolipid antigens and cancer therapy. Chemistry and Biology (1997). doi:10.1016/S1074-5521(97)90253-2

121. Birklé, S., Zeng, G., Gao, L., Yu, R. K. & Aubry, J. Role of tumor-associated gangliosides in cancer progression. Biochimie (2003). doi:10.1016/S0300-9084(03)00006-3

122. Ohtsubo, K. & Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. Cell (2006). doi:10.1016/j.cell.2006.08.019

123. Hakomori, S. Glycosylation defining cancer malignancy: New wine in an old bottle. Proceedings of the National Academy of Sciences of the United States of America (2002). doi:10.1073/pnas.172380699

124. Ohtsubo, K. et al. Dietary and genetic control of glucose transporter 2 glycosylation promotes insulin secretion in suppressing diabetes. Cell (2005). doi:10.1016/j.cell.2005.09.041

125. Collins, B. E. et al. Masking of CD22 by cis ligands does not prevent redistribution of CD22 to sites of cell contact. Proc. Natl. Acad. Sci. U. S. A. (2004). doi:10.1073/pnas.0400851101

126. Sharon, N. & Lis, H. History of lectins: From hemagglutinins to biological recognition molecules. Glycobiology (2004). doi:10.1093/glycob/cwh122

127. Weinhold, B. et al. Genetic ablation of polysialic acid causes severe neurodevelopmental defects rescued by deletion of the neural cell adhesion molecule. J. Biol. Chem. (2005). doi:10.1074/jbc.M511097200

128. Angata, K. et al. Sialyltransferase ST8Sia-II assembles a subset of polysialic acid that directs hippocampal axonal targeting and promotes fear behavior. J. Biol. Chem. (2004). doi:10.1074/jbc.M403429200

129. Ornitz, D. M. et al. Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cells. Mol. Cell. Biol. (1992). doi:10.1128/mcb.12.1.240

130. Ai, X. et al. QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. J. Cell Biol. (2003). doi:10.1083/jcb.200212083

131. Miljan, E. A. & Bremer, E. G. Regulation of growth factor receptors by gangliosides. Science's STKE : signal transduction knowledge environment (2002). doi:10.1126/stke.2002.160.re15

132. Belenkaya, T. Y. et al. Drosophila Dpp morphogen movement is independent of dynamin-mediated endocytosis but regulated by the glypican members of heparan sulfate proteoglycans. Cell (2004). doi:10.1016/j.cell.2004.09.031

133. Partridge, E.A. et al. Regulation of cytokine receptors by golgi N-glycan processing and endocytosis. Science (80-.). (2004). doi:10.1126/science.1102109

134. Smith, K. D., Davies, M. J., Bailey, D., Renouf, D. V. & Hounsell, E. F. Analysis of the glycosylation patterns of the extracellular domain of the epidermal growth factor receptor expressed in Chinese hamster ovary fibroblasts. Growth Factors (1996). doi:10.3109/08977199609034572

135. Zhen, Y., Caprioli, R. M. & Staros, J. V. Characterization of glycosylation sites of the epidermal growth factor receptor. Biochemistry (2003). doi:10.1021/bi027101p

136. Stroop, C. J. M. et al. Characterization of the carbohydrate chains of the secreted form of the human epidermal growth factor receptor. Glycobiology (2000). doi:10.1093/glycob/10.9.901

137. Liu, Y. C. et al. Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. Proc. Natl. Acad. Sci. U. S. A. (2011). doi:10.1073/pnas.1107385108

138. Kaszuba, K. et al. N-Glycosylation as determinant of epidermal growth factor receptor conformation in membranes. Proc. Natl. Acad. Sci. U. S. A. (2015).

doi:10.1073/pnas.1503262112

139. Taylor, E. S., Pol-Fachin, L., Lins, R. D. & Lower, S. K. Conformational stability of the epidermal growth factor (EGF) receptor as influenced by glycosylation, dimerization and EGF hormone binding. Proteins Struct. Funct. Bioinforma. (2017). doi:10.1002/prot.25220

140. Azimzadeh Irani, M., Kannan, S. & Verma, C. Role of N-glycosylation in EGFR ectodomain ligand binding. Proteins Struct. Funct. Bioinforma. (2017). doi:10.1002/prot.25314

141. Karaçali, S., □zzeto□lu, S. & Deveci, R. Glycosylation changes leading to the increase in size on the common core of N-glycans, required enzymes, and related cancer-associated proteins. Turkish Journal of Biology (2014). doi:10.3906/biy-1406-94

142. Pinho, S. S. & Reis, C. A. Glycosylation in cancer: Mechanisms and clinical implications. Nature Reviews Cancer (2015). doi:10.1038/nrc3982

143. Narasimhan, S. et al. DNP-Supported Solid-State NMR Spectroscopy of Proteins Inside Mammalian Cells. Angew. Chemie Int. Ed. (2019). doi:10.1002/anie.201908849

144. Hartmann, S. R. & Hahn, E. L. Nuclear double resonance in the rotating frame. Phys. Rev. (1962). doi:10.1103/PhysRev.128.2042

145. Pines, A., Gibby, M. G. & Waugh, J. S. Proton-enhanced NMR of dilute spins in solids. J. Chem. Phys. (1973). doi:10.1063/1.1680061

146. Reif, B. & Griffin, R. G. 1H detected 1H,15N correlation spectroscopy in rotating solids. J. Magn. Reson. (2003). doi:10.1016/S1090-7807(02)00035-6

147. Chevelkov, V. et al. 1H detection in MAS solid-state NMR spectroscopy of biomacromolecules employing pulsed field gradients for residual solvent suppression. J. Am. Chem. Soc. (2003). doi:10.1021/ja029354b

148. Andreas, L. B., Le Marchand, T., Jaudzems, K. & Pintacuda, G. High-resolution proton-detected NMR of proteins at very fast MAS. J. Magn. Reson. (2015). doi:10.1016/j.jmr.2015.01.003

149. Medeiros-Silva, J. et al. 1H-Detected Solid-State NMR Studies of Water-Inaccessible Proteins In Vitro and In Situ. Angew. Chemie - Int. Ed. (2016). doi:10.1002/anie.201606594

150. Hall, D. A. et al. Polarization-enhanced NMR spectroscopy of biomolecules in frozen solution. Science (80-.). (1997). doi:10.1126/science.276.5314.930

151. Carver, T. R. & Slichter, C. P. Polarization of nuclear spins in metals [15]. Physical Review (1953). doi:10.1103/PhysRev.92.212.2

152. Sauvée, C. et al. Highly efficient, water-soluble polarizing agents for dynamic nuclear polarization at high frequency. Angew. Chemie - Int. Ed. (2013). doi:10.1002/anie.201304657

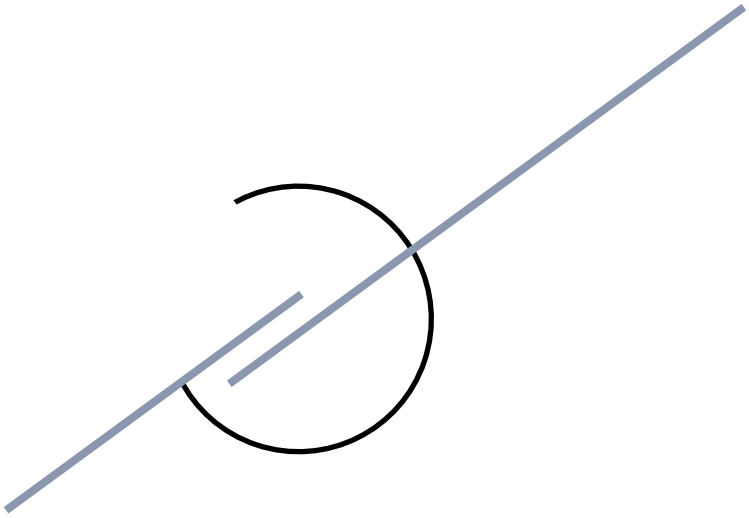
153. Lilly Thankamony, A. S., Wittmann, J. J., Kaushik, M. & Corzilius, B. Dynamic nuclear polarization for sensitivity enhancement in modern solid-state NMR. Progress in Nuclear Magnetic Resonance Spectroscopy (2017). doi:10.1016/j.pnmrs.2017.06.002

154. Baldus, M. & Meier, B. H. Total correlation spectroscopy in the solid state. the use of scalar couplings to determine the through-bond connectivity. Journal of Magnetic Resonance - Series A (1996). doi:10.1006/jmra.1996.0137

155. Szeverenyi, N. M., Sullivan, M. J. & Maciel, G. E. Observation of spin exchange by two-dimensional fourier transform 13C cross polarization-magic-angle spinning. J. Magn. Reson. (1982). doi:10.1016/0022-2364(82)90213-X

156. Hohwy, M., Rienstra, C. M., Jaroniec, C. P. & Griffin, R. G. Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy. J. Chem. Phys. (1999). doi:10.1063/1.478702

157. Schneider, R. et al. Probing molecular motion by double-quantum (13C,13C) solid-state NMR spectroscopy: Application to ubiquitin. J. Am. Chem. Soc. (2010). doi:10.1021/ja906283h



Chapter 2

Cellular solid-state NMR Spectroscopy

Reinier Damman, Siddarth Narasimhan,

Markus Weingarth and Marc Baldus

NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands

Chapter 9 Cellular Solid-state NMR Spectroscopy. in In-cell NMR Spectroscopy: From Molecular Sciences to Cell Biology 131–151 (The Royal Society of Chemistry, 2020)

ABSTRACT

Increasing evidence suggests that the highly complex and dynamic environment of bacterial and human cells imposes critical control on cellular functions which are difficult to mimic under in vitro conditions. Complementary to high-resolution light microscopy and electron tomography, in-cell solution-state NMR can track such structural and dynamical interactions at the most detailed, i.e., atomic level, provided that the molecular units tumble rapidly. Cellular solid-state NMR (ssNMR), on the other hand, provides increasing possibilities to probe molecular structure in bacterial and human cell preparations largely irrespective of molecular size. Here we review preparative as well as methodological aspects of such studies and highlight recent applications focussing on proteins embedded in their natural cell membrane environment.

INTRODUCTION

Structural biology has revolutionized our view of molecular life by providing three-dimensional views of molecular structures largely obtained from in-vitro preparations. Increasing efforts in the field of cellular structural biology are devoted to obtaining and refining such data under physiological conditions with the ultimate aim to describe cellular life and disease *in situ*. Solution-state (in-cell) NMR can track such structural and dynamical interactions in living cells at the most detailed, i.e., atomic level, provided that the molecular units are small and tumble rapidly. Impressive progress has been made to study protein structure, dynamics and biological processes under such conditions (see, e.g., Ref.[1-5] for recent reviews). In the following we discuss methods and applications underlining the growing potential of solid-state NMR (ssNMR) to provide complementary insight into complex cellular systems. Starting more than 40 years ago, the use of ssNMR spectroscopy for elucidating complex biomolecular systems such as collagen[6], bone[7], natural bilayers[8], bacterial cell walls[9], and cellular organelles[10] has been demonstrated. In many cases, such studies involved the use of one-dimensional ssNMR experiments which to this date[11] have provided valuable global and, in selected cases, molecule-specific insight into the structural organization of cells and their response to external stimuli. On the other hand, strong progress has been made in the last 15 years to develop and apply multi-dimensional ssNMR methodology for comprehensive determination of 3D (supra)molecular structure and dynamics of proteins and other biomolecules[12-15]. Combination of such methods with high sensitivity methods, in particular Dynamic Nuclear Polarization (DNP, see e.g. ref.[16]), as well as ^{1}H detected ssNMR (see e.g., ref.[17,18]) and advanced biochemical and cell biology approaches offers entirely new avenues to study molecular structure and organization in a cellular environment. In the following, we introduce such approaches and discuss their use for probing protein structure in bacterial cells and at the cell membrane periphery of human cells. Finally, we also give an outlook to future areas of research and application of cellular solid-state NMR spectroscopy.

EXPERIMENTAL

Sample preparation and isotope labelling for bacterial and mammalian cells

To conduct molecule-specific studies in a cellular context, isotope labeling is imperative to enhance NMR signals and to allow for multi-dimensional ssNMR correlation spectroscopy. Today's flexibility of cellular expression systems has enabled production of isotope enriched (prokaryotic and eukaryotic) proteins for in vitro studies using NMR. The transition from in vitro to in situ studies is however not straightforward and so far has required the development of dedicated protocols. The fundamental challenge for in situ NMR experiments lies in obtaining NMR signals of the protein of interest in a complex molecular background with sufficient spectroscopic sensitivity[1-5]. Our group has been investigating such molecule-specific labelling and sample preparation strategies for ssNMR wherein a target protein of interest (often overexpressed) is the only or most abundantly labelled moiety in the sample. A particularly relevant area of research for such preparations are membrane proteins where the surrounding lipid bilayer is critical for protein structure and function[19-22]. In such cases, ssNMR has, for example, shown an exceptional ability to study membrane proteins[14,15,23,24] including the examination of ligand binding[25-27], complex formation[28-30] and the influence of the surrounding lipid bilayer protein membrane protein structure and dynamics[31,32]. For

cellular ssNMR studies of membrane proteins, the isolation of the protein of interest in its native membrane environment is critical. In particular, procedures to be used must ensure that the membrane lipid composition which can influence membrane protein structure and function[20-22] is preserved. In bacteria, molecule-specific labelling can be performed by silencing the native polymerases using antibiotics including rifampicin (Figure 1A), provided that a non-bacterial polymerase such as T7 RNA polymerase is used for expression[33].

If obtaining such a construct is challenging, deletion methods in which endogenously highly expressed proteins such as OMP A/F present in the bacterial outer membrane are targeted can be applied[34]. Both type of preparations can be readily combined with techniques that contain the membrane protein of interest in bacterial cell envelopes or inner or outer membrane fractions[34-38]. For soluble prokaryotic proteins, overexpression can be used[39] but care must be exercised including the use of low temperatures[39,40] when conducting ssNMR experiments on whole prokaryotic or eukaryotic cells where temperature and mechanical forces as a result of magic angle spinning may impact cell intactness. For the purpose of studying mammalian cell cultures, different preparative methods have been used and can be subdivided into two main classes. The first class relies on the autonomous expression of the protein of interest in the host of interest. Banci et al. have used transient transfection of host cells which were grown in a growth medium containing labelled isotopes (see, e.g. Ref.[4] for a recent review). By using a strong promoter for a short amount of time, large quantities of the protein of interest are produced, ensuring high degree of labelling for this protein and reducing background labelling[4]. Alternatively, one can also use mammalian immortalized cell lines which are known for a high endogenous overexpression of specific proteins. For example, we previously utilized A431 cells, which express high amounts of the Epidermal Growth Factor Receptor (EGFR)[41]. Universal labelling of the A431 cells and vesicles (Figure 1B) enabled studies on the overall motion of EGFR, which can be supported by specific labelling schemes of sequential amino acid pairs to obtain information on local structure, motion and ligand binding[41] (*vide infra*).

The second class of preparative approaches introduces labelled biomolecules into an unlabelled host cell using dedicated protein delivery methods (Figure 1C). For example, the protein of interest can be fused to a peptide sequence which will allow for the translocation across the plasma membrane. As an example, Ubiquitin was shown to be translocated by such an approach across the plasma membrane where it interacts with the cellular environment[42]. In addition, pore-forming toxins have been used for protein delivery into human cells[43] and micro-injection of labelled material into large host cells, e.g. xenopus oocytes, has been demonstrated[1]. More recently, electroporation (Figure 1C) has been used for protein delivery into cells. During the electroporation procedure, small holes temporarily appear in the plasma membrane which allows for labelled material to diffuse into the cell. For example, such procedures have been applied to HEK293 or HELA cells[44] and recent work in our laboratory suggests that electroporation is also compatible with ssNMR studies (Narasimhan et al., in revision). In general, the benefit of using unlabelled host cells relates to the reduction of background labelling of other proteins or metabolites that may otherwise dominate NMR spectra. Moreover, the amount of labelled protein material can be readily controlled and the sample can be prepared homogenously. On the other hand, such procedures often inherently introduce proteins from a different organism than the studied host cell, an aspect that is not relevant when using an autonomous expression system.

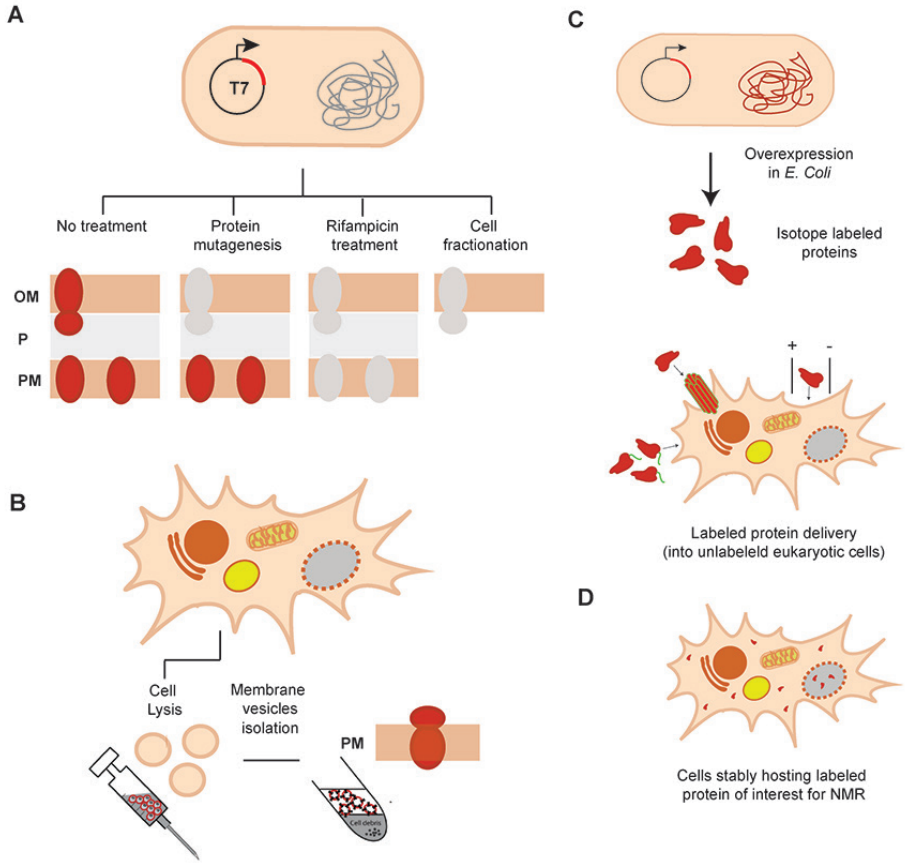


Figure 1 | Sample preparation strategies for ssNMR-based studies of prokaryotic and eukaryotic cellular systems. (A) Isotope-labeled membrane proteins can be produced in Escherichia coli using the bacteriophage T7 promoter-driven overexpression. The quality of cellular ssNMR spectra of cellular envelope (CE) and whole cell (WC) preparations can be significantly improved by the deletion of abundant *E. coli* proteins (OmpA and OmpF), rifampicin treatment during protein production and the isolation of the subcellular compartment of interest after cell fractionation. (B) Preparation of protein-rich membrane vesicles from human cells cultured in suitable growth media. (C) Protocol to study isotope-labeled soluble proteins produced in bacteria after using delivery methods in mammalian cells. (D) Transient transfection of host cells can be used to directly study soluble proteins in mammalian cells.

Dynamic Nuclear Polarization (DNP)

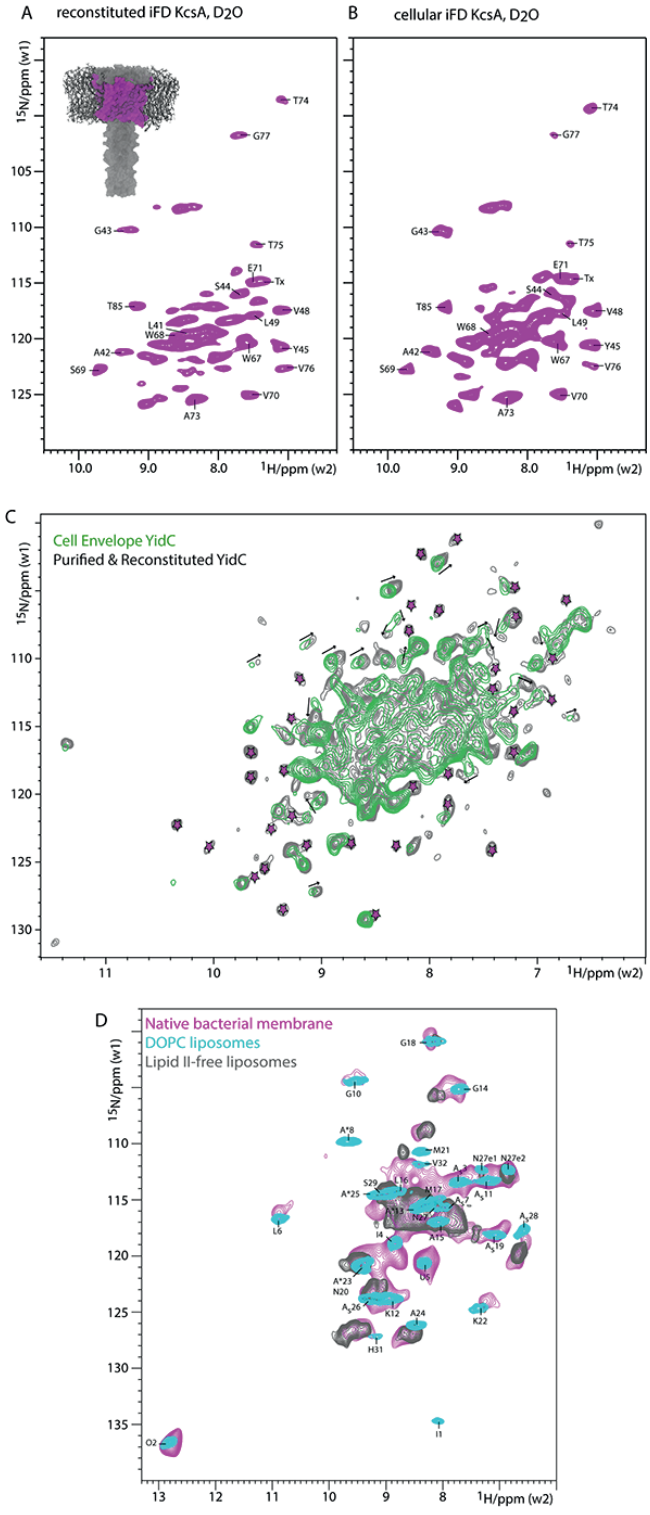
Spectroscopic sensitivity is of outermost importance for studying complex molecular systems and the advent of high-frequency Dynamic Nuclear Polarization (DNP[16,45]) spectrometers has been critical to research described in this review and has become the most widely applicable hyperpolarization technique to enhance the sensitivity of ssNMR experiments. Briefly, DNP takes advantage of the higher equilibrium polarization of electrons and enhances NMR signals by transferring the large polarization of electrons to nearby nuclei via microwave (μw)

irradiation. Usually, DNP studies involve the addition of soluble paramagnetic compounds to a sample dissolved in a glassy matrix at low temperatures and under magic-angle-spinning (MAS) conditions. Nuclear polarization is most efficiently generated by so-called cross effect DNP, involving a pair of electrons in the form of a biradical and a nuclear spin (See ref. [16]).

In 2012, we investigated the use of DNP to conduct ssNMR studies on ^{13}C,^{15}N-labeled cellular preparations of *E. coli* containing a recombinant membrane protein[40]. Bacterial cells and cell envelope extracts were typically suspended in a matrix consisting of a well-defined aqueous buffer containing d8-glycerol as cryoprotectant, and TOTAPOL[46] as polarizing agent. The biradicals TOTAPOL and AMUpol[47] are compatible with cell viability when used at a concentration of 5-40 mM and provide significant sensitivity enhancements at low temperature, i.e. at 100K. Regarding the maximum achievable DNP enhancement, recent work have shown that they can significantly vary, depending on sample preparation[48-50] and on the choice of the polarizing agent[51]. The enhancement generally reduces with increasing magnetic field in the case of biradicals that utilize the cross effect mediated by dipolar electron-electron interactions. Mance et al. have examined[52,53] the magnetic field dependence of biradicals including AMUPol. These studies suggested an approximately $1/B_0^2$ dependence of cross-effect DNP that is governed by a relative weak effective hyperfine coupling that describes the initial electron-nuclear polarization transfer step. The corresponding distance amounts to approximately 10 Ångstrom, which is in line with experimentally observed paramagnetic relaxation enhancement (PRE) effects under low temperature DNP conditions[54]. More recent results suggest that two-electron systems that exhibit sizable electron-electron exchange interactions exhibit a reduced magnetic field dependence and show maximum DNP enhancements at high field[51].

Next to sensitivity, spectroscopic resolution is another critical factor for DNP applications on complex (bio)molecules. In our previous work we have shown that high-field DNP can significantly enhance the prospects to conduct in-depth structural investigations of complex molecules such as membrane proteins. For example, we examined spectral resolution by comparing two-dimensional (^{13}C,^{13}C) and (^{15}N,^{13}C) correlation data sets on fully [^{13}C, ^{15}N]-labeled variants of the KcsA potassium channel at 800 MHz DNP conditions using soluble[49] and tagged[54] AMUPol variants to data obtained at ambient temperatures. In general, we observed that high field conditions can significantly enhance spectral resolution[49,55] and we also showed that the intrinsic paramagnetic properties of the polarizing agents can be used as direct structural probes during the spectroscopic analysis. A more detailed discussion of DNP in the context of cellular solid-state NMR can be found elsewhere[50].

Figure 2 | 1H detected dipolar-based 2D NH spectra of water-inaccessible regions in the potassium channel KcsA, acquired with A) reconstituted KcsA and B) directly in native cellular *E. coli* membranes. All data were recorded at 60 kHz MAS and 800 MHz. C) Overlay of dipolar-based ^{1}H-detected 2D NH spectra of YidC in cell envelopes (green) and purified and reconstituted in *E. coli* lipids (gray) at 55 kHz MAS and 800 MHz. Although overall the spectra agree well (>30 resolved peaks remain unchanged), 18 resolved peaks move (arrows) and >35 intensities change (stars) in the purified sample, relative to the cell envelope, suggesting that the structure and likely dynamics of YidC are different in native membranes. Figure adapted from Structure, 26, Baker, T. et al., 161-170, copyright (2018) with permission from Elsevier. D) Comparison of ^{1}H-detected 2D NH spectra of lipid II-bound nisin in native M. flavus membranes (magenta) and in DOPC (cyan). The gray spectrum shows nisin non-specifically bound to DOPG:DOPC liposomes (7:3 ratio) in the absence of lipid II. ➤


A
reconstituted iFD KcsA, D2O
B
cellular iFD KcsA, D2O
15N/ppm (w1)
1H/ppm (w2)
T74
G77
G43
T75
E71
Tx
S44
T85
V48
L41
W68
L49
A42
W67
Y45
S69
V76
V70
A73
C
Cell Envelope YidC
Purified & Reconstituted YidC
D
Native bacterial membrane
DOPC liposomes
Lipid II-free liposomes

Proton-detected ssNMR

While biological ssNMR historically used heteronuclear (low-gamma) detection, the emergence of dedicated deuteration schemes[36,56,57] and the commercial introduction of very high MAS probeheads[58] increasingly enables the direct detection of proton signals in complex biomolecular systems. ^{1}H-detection can provide substantial sensitivity gains, which is one of the reasons why it is commonly used in solution NMR studies, and also much reduces the necessary amount of sample. While the theoretically achievable spectral sensitivity in ^{1}H-detected experiments is much smaller compared to the achievable DNP signal enhancement, ^{1}H-detection is nevertheless an attractive alternative for cellular studies because of the usually markedly better spectral resolution and the possibility to conduct experiments at truly physiological temperatures and buffer conditions. In a first study, Weingarth et al. demonstrated[36] the use of ^{1}H-detection to quantitatively study membrane proteins directly in native cellular membranes (Figure 2A, B).

Here, we overexpressed the bacterial potassium channel KcsA in the *E. coli* inner membrane while the endogenous protein background was silenced by adding rifampicin to the ^{13}C,^{15}N-labelled growth medium[33]. Remarkably, ^{1}H-detected spectra in cellular membranes exhibit basically the same favorable resolution as spectra in synthetic liposomes, and readily enabled the annotation of a larger number of signals, and the analysis of protein dynamics under cellular conditions. More recently, we integrated this cellular ^{1}H-detected ssNMR approach with electron cryotomography to study YidC (Figure 2C), a bacterial membrane protein involved in the insertion of proteins into the inner membrane, in its native membrane environment[37]. These two methods are ideal complements and provide a powerful synergetic perspective to acquire structural data across several time- and length scales. In the case of YidC, we obtained a remarkably well resolved spectrum in the native cellular membrane, and this spectrum showed clear changes relative to synthetic membranes, implying structural and/or motional differences in the native environment. Finally, we recently demonstrated a hybrid approach based on ^{1}H-detection and high-field (800 MHz) DNP to study antibiotics directly in native cellular membranes[59] (Figure 2D). Here, we investigated the binding mode of the antimicrobial compound nisin to the so-called bacterial 'Achilles Heel' Lipid II, embedded in the native plasma membrane of *Micrococcus flavus*. We observed clear differences in native membranes, which provides precious information on pharmaceutical hot spots that can be used to rationally improve drug activity. Notably, we succeeded to perform this study with a minute amount of ~5 nmol of ^{13}C,^{15}N-labelled antibiotic, which was critical to zoom in on specific binding events.

Cellular ssNMR and complementary methods

During various stages of cellular ssNMR studies, a detailed characterization of cellular preparations by complementary methods is crucial. As discussed above, the concentration of labeled target protein must be high enough compared to other proteins to provide well-resolved resonances and unambiguous signal identification by ssNMR. Correspondingly, studies of protein composition and quantification are vital in the early stages of cellular ssNMR studies. The protein composition of cellular preparations can be readily investigated in a semi-quantitative manner by polyacrylamide gel electrophoresis (PAGE). Gel Glycine–SDS-PAGE are the commonly used SDS electrophoretic techniques for separating proteins in a mass range of 1-500 kDa. When protein identification becomes challenging in one dimension, 2-D electrophoretic systems in conjunction with protein identification techniques can be

used to establish proteome maps and yield information about the occurrence and abundance of proteins in cellular preparations. An example, of such studies is show in Figure 3A, where (A, lane 1) results of a Tris/Glycine SDS-PAGE of cell envelope of BL21 WT cells expressing the core complex of the type 4 secretion system (T4SScc, vide infra) that comprises three proteins (VirB7,9 and 10) are shown. In this case, endogenously highly expressed outer membrane proteins OmpF and OmpA are readily visible. Their contributions can be removed by using cell envelope preparations of BL21dm cells in which both proteins are deleted (lanes 2 and 3).

Alternatively, protein expression levels can be examined by mass spectrometry (MS). For example, we conducted MS experiments[41] on A431 cells and the isolated membrane vesicles probe the level of EGFR expression (see section 3.3). When analyzing whole cells or cellular preparations encompassing several cell compartments, the homogeneous targeting of the protein of interest represents another critical factor for obtaining high-quality ssNMR spectra and physiologically relevant information. Fluorescence microscopy in combination with fluorescently active ligands is an extremely powerful tool to localize the protein of interest. For example, the use of confocal microscopy on A431 cells and EGFR negative cells confirmed that EGFR is fully targeted to the membrane. For further characterization and subsequent structural analysis, FRET experiments as well as cryo-electron microscopy[35,41] and tomography[37] can be used in conjunction with cellular solid-state NMR studies. For example, we performed single molecular cryo-EM studies on type-4 secretion core complex (T4SScc) in cellular ssNMR preparations and after subsequent solubilisation[35] (Figure 3B). More recently, we combined ^{1}H ssNMR with cryo-EM tomography[37] to examine cell envelope preparations of bacterial cells overexpressing YidC in the inner membrane (Figure 3C). With the increasing resolution power of EM and ET methods, combined studies involving cellular ssNMR offer exciting opportunities to study proteins and other biomolecules in cells at highest spatial resolution. In addition in-cell solution-state NMR, EPR[44,60] can provide insights on the protein conformation in cellular settings. Due the fast relaxation of nitroxide spin labels at room temperature in general favors EPR experiments to be conducted at cryogenic temperatures similar to DNP ssNMR conditions.

RESULTS AND DISCUSSION

Bacterial cell envelopes

The bacterial cell envelope is a complex multilayered structure. Gram-negative bacteria are surrounded by a thin peptidoglycan (PG) cell wall, which itself is surrounded by an outer membrane (OM) containing proteins, phospholipids and lipopolysaccharides (LPS) (Figure 4.A). Reliable protocols and a number of successful applications have emerged in the last years, allowing bacterial membrane proteins to be investigated in situ with atom-level resolution. These protocols involved β-barrel proteins inserted into the bacterial outer membrane (OMPs)[40], proteins embedded in the inner membrane of bacteria such as ion and proton channels[36,61], retinal proteins[22] and electron transport proteins[62] or periplasmic proteins associated with the peptidoglycan layer[63].

Previously, we showed[34] that prokaryotic cell envelope containing the intact inner membrane (IM) and the outer membrane (OM) may be obtained through mild and non-disruptive lysis procedure such as in a French press, followed by successive removal of unbroken cells

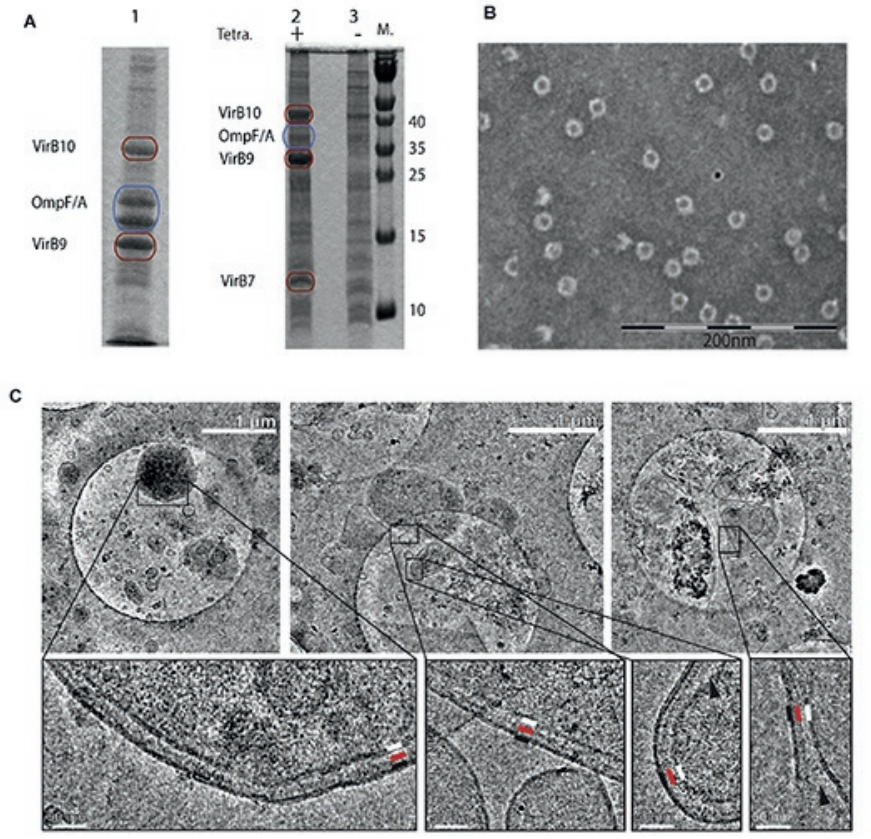


Figure 3 | (A) Lane 1: 10% Tris/Glycine SDS-PAGE of cell envelope of BL21 WT expressing T4SScc. Lanes 2 and 3: 17.5% Tris/Tricine SDS-PAGE of cell envelope of BL21dm expressing T4SScc (lane 2) and without expressing T4SScc (lane 3). M refers to the molecular weight marker. B) Electron Microscopy data of T4SScc after purification. C) The Morphology of Cell Envelope Samples over YidC overexpressing bacteria by cryoEM. (top) Cell envelopes were vitrified on graphene oxide-treated holey carbon grids. The membranes adopt a variety of shapes and sizes, with many maintaining a native ultrastructure. (bottom) Membrane structures of different shapes and sizes maintain the native cell wall architecture when viewed at higher magnification of outer membrane (black line), cell wall (red line), and inner membrane (white line). When the inner and outer membranes separate (black arrowheads), the cell wall appears to favour the outer membranes. Figure adapted from Structure, 26, Baker, T. et al., 161-170, copyright (2018) with permission from Elsevier.

and cytosolic components through different centrifugal steps. More recently[37], we could also confirm that IM and OM remain in such cell envelope (CE) preparations. The cell envelope approach firstly was applied on PagL, a 150-residue integral membrane-protein PagL from *Pseudomonas aeruginosa*, an OM enzyme that removes a fatty acyl chain from LPS and helps the bacterium to evade the host immune system. Using uniformly ^{13}C,^{15}N-labeled preparations of whole cells (WC) and cell envelopes (CE), we performed a set of 2D ^{13}C-^{13}C correlation experiments employing dipolar-based magnetization transfer steps. Both preparations yielded NMR spectra of astonishing quality considering sample complexity and

non-crystallinity, with well-dispersed cross-peaks characteristic for protein and lipid signals. As anticipated, we observed an improvement in both sensitivity and spectral resolution for the CE preparation, potentially due to the single contribution of CE-associated components. To examine in further detail the conformation of PagL in CE, we performed 2D ^{15}N-edited ^{13}C correlation experiments[33] in which signals arising from non-proteinaceous molecular components are largely reduced. Because of the favorable spectroscopic dispersion among Thr, Ser, and Gly residues in 2D and 3D CC/NC correlation experiments, we could subsequently perform a residue-specific analysis based on sequential resonance assignments of PagL residues located in different topological regions of the protein. Complementary protocols were further established to characterize major endogenous macromolecular components of *E. coli* CE that surround the target protein PagL such as the lipoprotein lpp or Braun's protein that belongs to the most abundant CE proteins in exponentially growing *E. coli* cells. Finally, highly-flexible non-proteinaceous CE-associated components like PG and LPS that play crucial roles in bacterial physiology, virulence and antibiotic resistance were readily identified using through-bond ^{1}H-^{13}C and ^{13}C-^{13}C correlation spectra obtained on CE and WC preparations (Figure 4.B). Using this strategy, we obtained *de novo* ssNMR assignments of LPS and PG glucosamine units and some of their substituents[34].

More recently we further extended our methodology to study protein machines spanning the entire bacterial cell envelope (Fig, 4,C). In particular we studied the structural organization of the type IV Secretion System Core Complex (T4SScc) consisting of 14 copies each of Vir- B7, B9 and B10 proteins, that spans the whole length of the periplasm in order to shuttle substrates/cargo in and out of the cells[35]. We used specific ([^{13}C,^{15}N]-(Gly,Ser,Leu,Val)) and sequentially selective ([^{13}C,^{15}N]-Thr, [^{15}N]-Val)) isotope labelling to reduce signal ambiguity and establish unique sequential correlations in cell envelope preparations of BL21dm cells. For our structural analysis, we exploited the well-established correlation between NMR resonance frequencies and protein secondary structure[64,65] to obtain structural information about the backbone fold of cell-embedded T4SScc using 3D and 2D ssNMR data sets (Figure 4C,). Comparison of our experimental results obtained under DNP conditions to *in silico* structural models helped to describe the fold of the T4SScc in the cellular envelope (Figure 4C), in particular of protein regions that have been elusive for X-ray or single particle cryo-EM studies. Moreover, the combination of DNP-ssNMR and ssNMR experiments conducted at ambient temperatures provided insight into the inner membrane embedded parts of VirB10.

Finally, cell envelope preparations can be further modified to isolate the inner or outer membrane fractions through density gradient centrifugation. Medeiros-Silva et al[36] could for the first time used a similar approach to study an ion-channel protein KcsA in the inner membrane of *E. coli.* Using proton detection, they reported significant differences in the protein structure in the native lipids as opposed to *in vitro* (Figure 2.A).

Intact bacteria cells

Similar to the use of ssNMR for cell envelope samples, experiments are also possible for intact bacterial cells provided that experimental conditions are used that keep cell morphology intact. In particular, these issues relate to centrifugal forces as well as temperature increase may affect bacterial cell structure. A convenient means to address these issues is the use of low temperature conditions that has been employed to study soluble proteins[39] and other biomolecules[40], possibly under DNP conditions[40]. In the latter case we could not only study

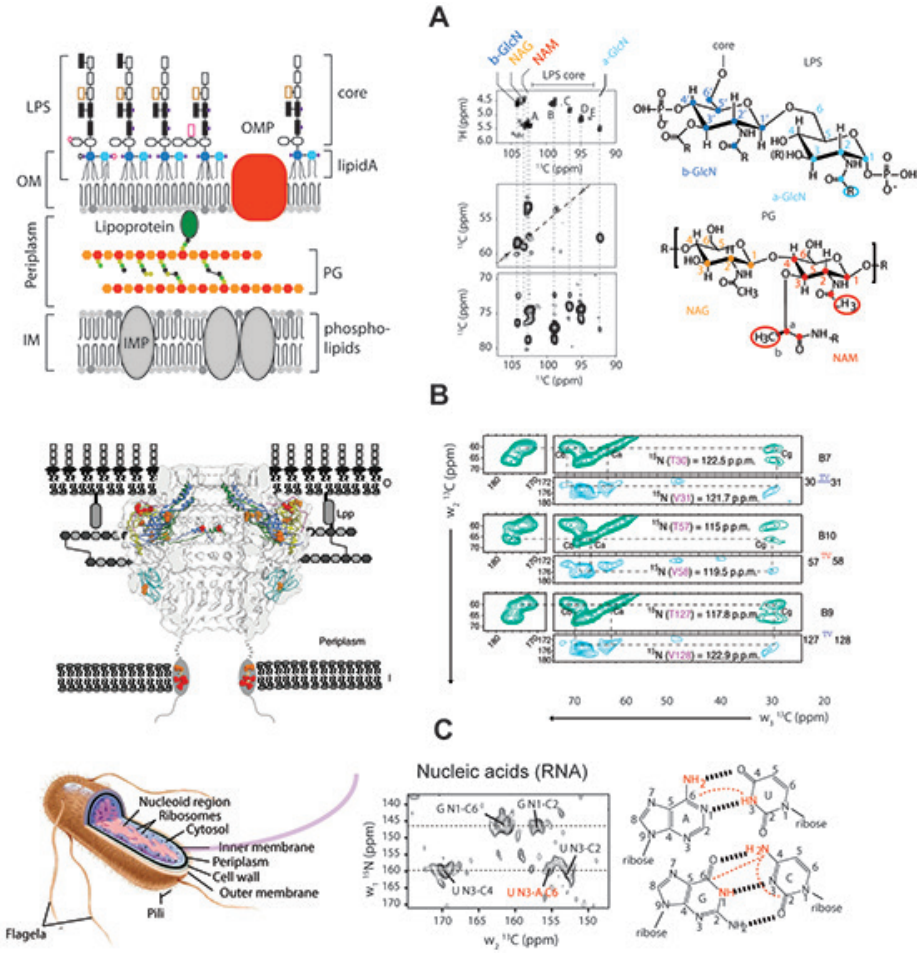


Figure 4 | (A, left) Overall structural organization of the cell envelope of E.coli including an outer membrane protein (OMP). (A, right) Resonance assignment of flexible LPS and PG moleculars using through-bond 2D (^{1}H-^{13}C) INEPT and 2D (^{13}C-^{13}C) INEPT-TOBSY correlation spectroscopy. The CC and HC correlations that belong to the same spin system are connected by dashed lines. Assigned resonances from LPS and PG glucosamine units are highlighted onto chemical structures. β/α-GlcN, β/ α-N-acetylglucosamine; PG NAG, Peptidoglycan N-acetyl Glucosamine; NAM, N-acetyl Muramic acid, LP, lipopolysaccharide. (B, right) Analysis of DNP-based 2D and 3D ssNMR data sets of tailored-labeled cell-embedded T4SSc. Left, 2D (^{13}C,^{13}C) planes taken out of 3D NCACX (green) and 3D NCOCX (light blue) of TV-T4SSc. (B, left), summary of residue-specific ssNMR probes and their location in reference to the cellular envelope including the T4SSc electron microscopy map with the outer membrane complex fitted inside (PDB: 3JQO). The N-terminus of VirB9 is also docked into the map. Identified residues in ssNMR spectra are shown as orange balls for the TV-T4SSc and red balls for the GSLV-T4SSc. C) Identification of nucleic acids in whole-cell (WC) ssNMR data of PagL overexpressing bacteria. Selected spectral regions from DNP-enhanced 2D NCO spectra obtained on (U-^{13}C,^{15}N)-labeled WC showing characteristic correlations of RNAs. Indicated assignments are based on BMRB average chemical shifts from common RNA nucleotides. Chemical structure of standard RNA bases with Watson-Crick base pairing throughout. Intra- and intermolecular contacts between cytosine N4 or uracil N3 and other C atoms that support our assignment are indicated.

proteinaceous species but our data also allowed to identify nucleic acids in selected spectral regions of 2D NCO spectra obtained on (U-^{13}C,^{15}N)-labeled whole cell samples showing characteristic correlations of RNAs (Figure 4D). Assignments indicated in this figure are based on BMRB average chemical shifts from common RNA nucleotides and Intra- and intermolecular contacts between cytosine N4 or uracil N3 and other C atoms that support our assignment are indicated. On-going work in our laboratory also suggests that more detailed studies of nucleotides and proteins and their complexes may be possible by carefully optimizing sample preparations including the use of protein delivery methods and the use of mixed labelling[66,67] to gain detailed structural and molecule-specific insight into the structural organization of complexes in cells.

Human cells

Using a DNP-supported cellular ssNMR approach and dedicated sample preparation methods (see above), we more recently studied structural and dynamical aspects of the activation mechanism of EGFR in membrane vesicles derived from human A431 cells (Figure 5A,B). We quantified EGFR expression levels using MS experiments[41] on A431 cells and the isolated membrane vesicles (Figure 5C-E). Using a well-established semi-quantitative analysis based on summed ions intensities over all detected peptides, we found actin to be the most abundant protein in whole A431 cells, followed by other abundant soluble molecules including heat shock and histone proteins. While the EGFR expression level was lower than these proteins, EGFR still represented the most abundant membrane protein in the cells, in line with previous findings. When moving to isolated membrane vesicles, we found EGFR relatively enriched by at least a factor 5.5, making EGFR, together with actin, the most abundant protein in the membrane vesicles (Figure 5D).

For our ssNMR analysis, we firstly probed rigid and dynamic regions of uniformly isotope-labeled receptor preparations by combing dipolar and scalar-based ssNMR experiments that can give complementary information about protein dynamics at ambient temperatures[41] (Figure 5F). In our studies we exploited the fact that intermediate-time scale motion will further influence dipolar recoupling efficiencies, for example when using dipolar-based Double Quantum-Single Quantum (DQSQ) correlation experiments[28,68.] When applied in 2D ssNMR spectroscopy, amino-acid specific information can be obtained when comparing data sets before and after receptor activation. Signal dispersion in such data sets arises from amino-acid type and protein secondary structure. For example, various crystal structures of EGFR domains predict that threonine and serine residues are more commonly found in α-helices in the kinase domain (KD), while they mostly reside in β-sheets or random coil protein regions in the extracellular domain (ECD). Using such dependencies, we could correlate ssNMR signal patterns before and after activation with the structure and intrinsic mobility of different EGFR domains (Figure 5G). In order to confirm the change in overall dynamics observed after activation of EGFR in the fully labelled samples, we designed local probes by specific amino acid labelling to study the local effect of receptor activation. Based on the amino acid sequence of EGFR and the other most abundant proteins in the plasma membrane preparation, two amino acids (^{13}C-Met, ^{13}C-Phe) were selected for ^{13}C-labelling, while two others (^{15}N-Thr, ^{15}N-Leu) were uniformly ^{15}N-labelled, creating unique signal sets when conducting sequential NCOCX-type of experiments[69,70]. Hence, these experiments allowed us to zoom in on specific protein segments know to be critical for receptor activation. We could confirm the EGF binding site and found that, upon binding of EGF

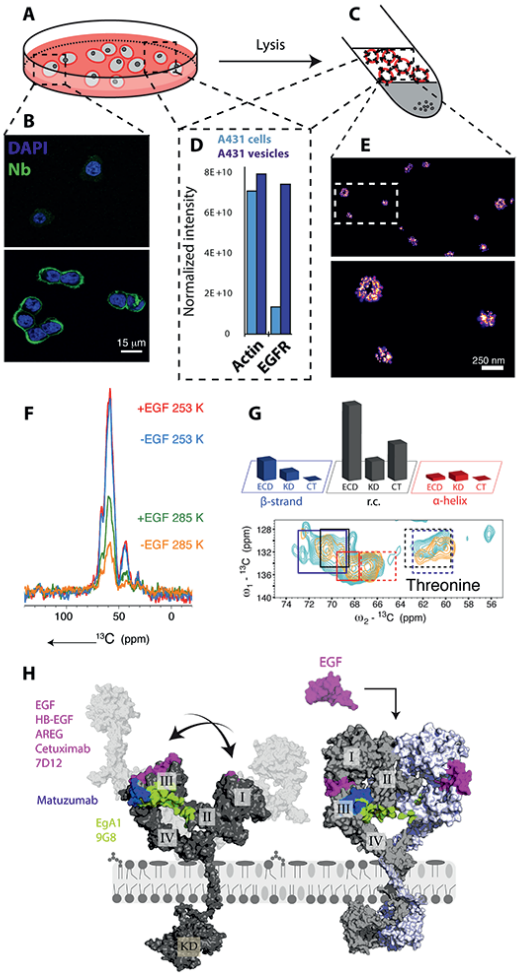


Figure 5 | Plasma membrane vesicle isolation allows for structural studies in native environments. A) A431 cells that over-express EGFR are cultured with (^{13}C,^{15}N)-labeled amino acids to label all proteins inside the cells. B) Confocal imaging of NIH 3T3 2.2 cells (EGFR negative) and A431 cells showing the nucleus by using DAPI (blue) and EGFR by using an anti-EGFR nanobody (green). C) A431 cells are harvested and lysed in order to form plasma membrane vesicles. D) Mass spectrometry analysis showing the most abundant proteins in the whole cell preparation and plasma membrane vesicle preparation. E) Super-resolution dSTORM imaging of the plasma membrane vesicles using nanobodies to stain for EGFR. F) First increment of a 2D (^{13}C-^{15}N)-NCα recorded on A431 plasma membrane vesicle in absence or presence of EGF at ambient or frozen temperature, highlighting the increase in rigidity in overall dynamics at ambient temperature after EGF addition. G) Zoom in on the threonine region in a (^{13}C-^{13}C) Double Quantum Single Quantum experiment recorded at ambient temperature in the absence (orange) or presence (blue) of EGF. Boxes indicate confidence intervals for secondary structure elements. Boxes for alpha-helical, random coil and β-sheet structures are coloured red, black and blue respectively. The dashed boxes indicate Cα resonances while the solid boxes indicate Cβ. The distribution of secondary structure elements per domain of EGFR is shown by bar the bar graphs for threonine. H) Model for EGFR activation, highlighting the dynamical ECD prior to EGF-induced activation leading the stabilization of the ECD and subsequent activation of the receptor. Several binding epitopes are indicated on EGFR of activating ligands (EGF, HB-EGF and AREG) and inactivating ligands (Cetuximab, 7D12, Matuzumab, EgA1 and 9G8).

to EGFR, spectral dispersion (i.e., secondary structure) is preserved and conformational heterogeneity is reduced. Taken together, the combination of ssNMR with DNP-ssNMR led us to model the activation step of EGFR that, according to our data, involves changes in overall as well as local receptor dynamics without significant changes in protein secondary structure (Figure 5H).

CONCLUSIONS

NMR is providing increasing opportunities to study protein structure under cellular conditions where a complex molecular environment as well as endogenously low protein concentrations require the combined use of state-of-the art protein chemistry, cell biology and NMR spectroscopic methods. In such studies, solid-state NMR can play an important role since molecular tumbling does not restrict its use, opening novel avenues to study structures and interactions of proteins and other biomolecules as well as the formation of biomolecular complexes in cells and cellular compartments. A critical parameter for such studies relates to spectroscopic sensitivity and significant progress has been made in developing ssNMR methods and protocols that make use of DNP and ^{1}H ssNMR for the study of cellular (so far mostly membrane-embedded) proteins. In our contribution, we described the conceptual details of such methods.

Thus far, most DNP-supported studies of complex biomolecules utilized chemical-shift information or relied on the determination of specific distance or torsional restraints. Similar to progress in the field of in-cell solution-state NMR[71,72] we envision that further efforts will be devoted in the next years to increase the number of structural constraints that can be obtained from such studies. For example, the rich arsenal of ^{13}C/^{15}N detected experiments could be tailored to give 3D structural information for cellular DNP-supported ssNMR conditions. Likewise, dedicated methods that probe the supramolecular setting[67] will be of high interest. Such studies may, at least for bacterial systems where deuteration is straightforward, also go hand-in-hand with cellular ssNMR ^{1}H detection approaches[26,36] and make use of MAS setups that go well beyond 60 kHz. As we have demonstrated recently[59,73], the combination of both high-sensitivity techniques may provide unique information on molecular structure and dynamics in a cellular context.

While the emphasis in many ssNMR studies so far has been largely been on membrane proteins, we[40] and others[39] have already shown that ssNMR studies can also be conducted on soluble molecules in whole cells. It seems likely that future cellular ssNMR studies may be possible on large cell-embedded complexes that are difficult to study by in-cell solution-state NMR. With these aspects in mind, cellular solid-state NMR may be able to provide detailed structural insight into biomolecules in their natural setting that is complementary to modalities such as cryo-electron tomography or high-resolution fluorescence spectroscopy. Applications may range from elucidating molecular mechanisms associated with bacterial pathogenicity and virulence to the examination of protein folding and aggregation in human cells with an important goal for determining drug modes of action and developing new-generation therapeutics in a cellular, i.e. physiological, context.

Acknowledgements
We gratefully acknowledge our collaborators and colleagues for their invaluable contributions to cited publications from our own research group. These studies were supported through grants from NWO, the EU and NIH.

REFERENCES

1 R. Hänsel, L. M. Luh, I. Corbeski, L. Trantirek and V. Dötsch, Angew. Chem. Int. Ed., 2014, 53, 10300–10314.

2 J. M. Plitzko, B. Schuler and P. Selenko, Curr. Opin. Struct. Biol., 2017, 46, 110–121.

3 A. Y. Maldonado, D. S. Burz and A. Shekhtman, Prog. Nucl. Magn. Reson. Spectrosc., 2011, 59, 197–212.

4 E. Luchinat and L. Banci, Acc. Chem. Res., 2018, 51, 1550–1557.

5 C. Li, J. Zhao, K. Cheng, Y. Ge, Q. Wu, Y. Ye, G. Xu, Z. Zhang, W. Zheng, X. Zhang, X. Zhou, G. Pielak and M. Liu, Annual Rev. Anal. Chem., 2017, 10, 157–182.

6 D. A. Torchia and D. L. Vanderhart, J Mol Biol, 1976, 104, 315–321.

7 Herzfeld, A. Roufosse, R. A. Haberkorn, R. G. Griffin and M. J. Glimcher, Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci., 1980, 289, 459–469.

8 S. O. Smith, I. Palings, V. Copie, D. P. Raleigh, J. Courtin, J. A. Pardoen, J. Lugtenburg, R. A. Mathies and R. G. Griffin, Biochemistry, 1987, 26, 1606–1611.

9 G. S. Jacob, J. Schaefer and G. E. Wilson, J. Biol. Chem., 1983, 258, 10824–10826.

10 A. C. Sivertsen, M. J. Bayro, M. Belenky, R. G. Griffin and J. Herzfeld, J Mol Biol, 2009, 387, 1032–1039.

11 J. A. H. Romaniuk and L. Cegelski, Philosophical Transactions of the Royal Society B: Biological Sciences, 2015, 370, 20150024–14.

12 R. G. Griffin, Nat. Struct. Biol., 1998, 5, 508–512.

13 S. Luca, H. Heise and M. Baldus, Acc. Chem. Res., 2003, 36, 858–865.

14 S. J. Ullrich and C. Glaubitz, Acc. Chem. Res., 2013, 46, 2164–2171.

15 S. Wang and V. Ladizhansky, Prog. Nucl. Magn. Reson. Spectrosc., 2014, 82, 1–26.

16 Q. Z. Ni, E. Daviso, T. V. Can, E. Markhasin, S. K. Jawla, T. M. Swager, R. J. Temkin, J. Herzfeld and R. G. Griffin, Acc. Chem. Res., 2013, 46, 1933–1941.

17 P. Fricke, V. Chevelkov, M. Zinke, K. Giller, S. Becker and A. Lange, Nat Protoc, 2017, 12, 764–782.

18 S. Asami and B. Reif, Acc. Chem. Res., 2013, 46, 2089–2097.

19 M. Hong, Y. Zhang and F. Hu, Annu. Rev. Phys. Chem., 2012, 63, 1–24.

20 H.-X. Zhou and T. A. Cross, Annu. Rev. Biophys., 2013, 42, 361–392.

21 E. A. W. van der Cruijsen, A. V. Prokofyev, O. Pongs and M. Baldus, Biophys J, 2017, 112, 99–108.

22 M. E. Ward, S. Wang, R. Munro, E. Ritz, I. Hung, P. L. Gor'kov, Y. Jiang, H. Liang, L. S. Brown and V. Ladizhansky, Biophysj, 2015, 108, 1683–1696.

23 Herzfeld and J. C. Lansing, Annu. Rev. Biophys. Biomolec. Struct., 2002, 31, 73–95.

24 M. Kaplan, C. Pinto, K. Houben and M. Baldus, Q. Rev. Biophys., 2016, 49, e15.

25 V. S. Bajaj, M. L. Mak-Jurkauskas, M. Belenky, J. Herzfeld and R. G. Griffin, Proc Natl Acad Sci USA, 2009, 106, 9244–9249.

26 A. Lange, K. Giller, S. Hornig, M.-F. Martin-Eauclaire, O. Pongs, S. Becker and M. Baldus, Nature, 2006, 440, 959–962.

27 L. Joedicke, J. Mao, G. Kuenze, C. Reinhart, T. Kalavacherla, H. R. A. Jonker, C. Richter, H. Schwalbe, J. Meiler, J. Preu, H. Michel and C. Glaubitz, Nat Chem Biol, 2018, 14, 284–290.

28 M. Etzkorn, K. Seidel, L. Li, S. Martell, M. Geyer, M. Engelhard and M. Baldus, Structure, 2010, 18, 293–300.

29 S. Wang, R. A. Munro, L. Shi, I. Kawamura, T. Okitsu, A. Wada, S.-Y. Kim, K.-H. Jung, L. S. Brown and V. Ladizhansky, Nat Meth, 2013, 10, 1007–1012.

30 J. Maciejko, M. Mehler, J. Kaur, T. Lieblein, N. Morgner, O. Ouari, P. Tordo, J. Becker-Baldus and C. Glaubitz, J Am Chem Soc, 2015, 137, 9032–9043.

31 C. Ader, O. Pongs, S. Becker and M. Baldus, Biochim. Biophys. Acta-Biomembr., 2010, 1798, 286–290.

32 D. Huster, K. Arnold and K. Gawrisch, J Phys Chem B, 1999, 103, 243–251.

33 L. A. Baker, M. Daniëls, E. A. W. van der Cruijsen, G. E. Folkers and M. Baldus, J Biomol NMR, 2015, 62, 199–208.

34 M. Renault, R. Tommassen-van Boxtel, M. P. Bos, J. A. Post, J. Tommassen and M. Baldus, Proceedings of the National Academy of Sciences, 2012, 109, 4863–4868.

35 M. Kaplan, A. Cukkemane, G. C. P. van Zundert, S. Narasimhan, M. Daniëls, D. Mance, G. Waksman, A. M. J. J. Bonvin, R. Fronzes, G. E. Folkers and M. Baldus, Nat Meth, 2015, 12, 649–652.

36 J. Medeiros Silva, D. Mance, M. Daniëls, S. Jekhmane, K. Houben, M. Baldus and M. Weingarth, Angew. Chem. Int. Ed., 2016, 55, 13606–13610.

37L. A. Baker, T. Sinnige, P. Schellenberger, J. de Keyzer, C. A. Siebert, A. J. M. Driessen, M. Baldus and K. Grünewald, Structure, 2018, 26, 161–170.e3.

38C. Pinto, D. Mance, M. Julien, M. Daniëls, M. Weingarth and M. Baldus, Journal of Structural Biology, 2017, –.

39S. Reckel, J. J. Lopez, F. Löhr, C. Glaubitz and V. Dötsch, Chembiochem, 2012, 13, 534–537.

40M. Renault, S. Pawsey, M. P. Bos, E. J. Koers, D. Nand, R. Tommassen-van Boxtel, M. Rosay, J. Tommassen, W. E. Maas and M. Baldus, Angew. Chem. Int. Ed., 2012, 51, 2998–3001.

41M. Kaplan, S. Narasimhan, C. de Heus, D. Mance, S. Van Doorn, K. Houben, D. Popov-Celeketic, R. Damman, E. A. Katrukha, P. Jain, W. J. C. Geerts, A. J. R. Heck, G. E. Folkers, L. C. Kapitein, S. Lemeer, P. M. P. V. B. E. Henegouwen and M. Baldus, Cell, 2016, 167, 1241.

42D. Sakakibara, A. Sasaki, T. Ikeya, J. Hamatsu, T. Hanashima, M. Mishima, M. Yoshimasu, N. Hayashi, T. Mikawa, M. Walchli, B. O. Smith, M. Shirakawa, P. Guntert and Y. Ito, Nature, 2009, 458, 102–U10.

43S. Ogino, S. Kubo, R. Umemoto, S. Huang, N. Nishida and I. Shimada, J Am Chem Soc, 2009, 131, 10834–10835.

44F.-X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H. M. Rose, M. Stuiver, S. Verzini, D. Lorenz, M. van Rossum, D. Goldfarb and P. Selenko, Nature, 2016, 530, 45–50.

45D. A. Hall, D. C. Maus, G. J. Gerfen, S. J. Inati, L. R. Becerra, F. W. Dahlquist and R. G. Griffin, Science, 1997, 276, 930–932.

46C. Song, K.-N. Hu, C.-G. Joo, T. M. Swager and R. G. Griffin, J Am Chem Soc, 2006, 128, 11385–11390.

47C. Sauvee, M. Rosay, G. Casano, F. Aussenac, R. T. Weber, O. Ouari and P. Tordo, Angew. Chem. Int. Ed., 2013, 52(41):10858-61.

48T. Jacso, W. T. Franks, H. Rose, U. Fink, J. Broecker, S. Keller, H. Oschkinat and B. Reif, Angewandte Chemie, 2012, 124, 447–450.

49E. J. Koers, E. A. W. van der Cruijsen, M. Rosay, M. Weingarth, A. Prokofyev, C. Sauvee, O. Ouari, J. van der Zwan, O. Pongs, P. Tordo, W. E. Maas and M. Baldus, J Biomol NMR, 2014, 60, 157–168.

50A. L. Paioni, M. A. M. Renault and M. Baldus, EMAGRES, 2018, {7}, {51–61}.

51G. Mathies, M. A. Caporini, V. K. Michaelis, Y. Liu, K.-N. Hu, D. Mance, J. L. Zweier, M. Rosay, M. Baldus and R. G. Griffin, Angew. Chem. Int. Ed., 2015, 54, 11770–11774.

52D. Mance, P. Gast, M. Huber, M. Baldus and K. L. Ivanov, J. Chem. Phys., 2015, 142, –.

53P. Gast, D. Mance, E. Zurlo, K. L. Ivanov, M. Baldus and M. Huber, Phys. Chem. Chem. Phys., 2017, 19, 3777–3781.

54E. A. W. van der Cruijsen, E. J. Koers, C. Sauvee, R. E. Hulse, M. Weingarth, O. Ouari, E. Perozo, P. Tordo and M. Baldus, Chem. Eur. J., 2015, 21, 12971–12977.

55P. Fricke, D. Mance, V. Chevelkov, K. Giller, S. Becker, M. Baldus and A. Lange, J Biomol NMR, 2016, 65, 121–126.

56V. Chevelkov, K. Rehbein, A. Diehl and B. Reif, Angew. Chem. Int. Ed., 2006, 45, 3878–3881.

57D. Mance, T. Sinnige, M. Kaplan, S. Narasimhan, M. Daniëls, K. Houben, M. Baldus and M. Weingarth, Angewandte Chemie, 2015, 127, 16025–16029.

58D. Lalli, M. N. Idso, L. B. Andreas, S. Hussain, N. Baxter, S. Han, B. F. Chmelka and G. Pintacuda, J Am Chem Soc, 2017, 139, 13006–13012.

59J. Medeiros Silva, S. Jekhmane, A. L. Paioni, K. Gawarecka, M. Baldus, E. Swiezewska, E. Breukink and M. Weingarth, Nature Communications, 2018, 9, 3963.

60A. Martorana, G. Bellapadrona, A. Feintuch, E. Di Gregorio, S. Aime and D. Goldfarb, J Am Chem Soc, 2014, 136, 13458–13465.

61Y. Miao, H. Qin, R. Fu, M. Sharma, T. V. Can, I. Hung, S. Luca, P. L. Gor'kov, W. W. Brey and T. A. Cross, Angew. Chem. Int. Ed., 2012, 51, 8383–8386.

62K. Yamamoto, M. A. Caporini, S.-C. Im, L. Waskell and A. Ramamoorthy, Biochim. Biophys. Acta, 2015, 1848, 342–349.

63P. Schanda, S. Triboulet, C. Laguri, C. M. Bougault, I. Ayala, M. Callon, M. Arthur and J.-P. Simorre, J Am Chem Soc, 2014, 136, 17852–17860.

64S. Neal, A. M. Nip, H. Y. Zhang and D. S. Wishart, J Biomol NMR, 2003, 26, 215–240.

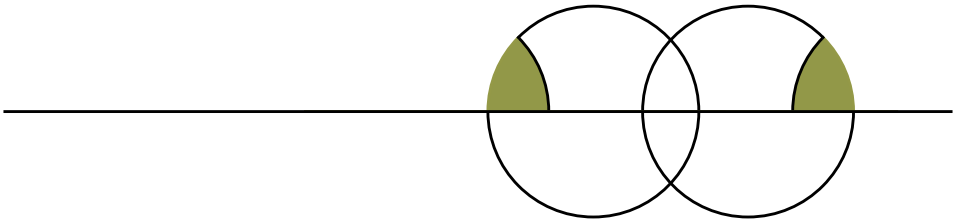
65K. Seidel, M. Etzkorn, R. Schneider, C. Ader and M. Baldus, Solid State Nucl Magn Reson, 2009, 35, 235–242.

66M. Etzkorn, A. Bockmann, A. Lange and M. Baldus, J Am Chem Soc, 2004, 126, 14746–14751.

67M. Weingarth and M. Baldus, Acc. Chem. Res., 2013, 46, 2037–2046.

68R. Schneider, K. Seidel, M. Etzkorn, A. Lange, S. Becker

and M. Baldus, J Am Chem Soc, 2010, 132, 223–233.
69 A. T. Petkova, M. Baldus, M. Belenky, M. Hong, R. G. Griffin and J. Herzfeld, J. Magn. Reson., 2003, 160, 1–12.
70 M. Baldus, Prog. Nucl. Magn. Reson. Spectrosc. 41, 1-47, 2002.
71 T. Ikeya, T. Hanashima, S. Hosoya, M. Shimazaki, S. Ikeda, M. Mishima, P. Güntert and Y. Ito, Sci. Rep., 2016, 1–11.
72 T. Müntener, D. Häussinger, P. Selenko and F.-X. Theillet, J. Phys. Chem. Lett., 2016, 7, 2821–2825.
73 C. Pinto, D. Mance, T. Sinnige, M. Daniëls, M. Weingarth and M. Baldus, Nature Communications, 2018, 9, 4135.



Chapter 3

Allosteric nanobodies reveal the importance of regulating conformational dynamics in EGFR functioning

Reinier Damman[1,2], Apfrida A. Kendek[1,2],

Paul M.P. van Bergen en Henegouwen[2] and Marc Baldus[1]

[1] NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht UniversityThe Netherlands.
[2] Cell Biology, Neurobiology and Biophysics, Department of Biology, Faculty of Science, Utrecht University, Utrecht, the Netherlands.

Manuscript in preparation

ABSTRACT

The epidermal growth factor receptor (EGFR) is a receptor tyrosine kinase which is crucial for the regulation of cell proliferation, migration and differentiation and is therefore a common target in anti-cancer therapy. By binding to the extracellular domain (ECD), ligands induce the stabilization of the open conformation, promoting receptor dimerization and oligomerization which results in kinase activation and receptor downregulation. The ECD samples a large conformational landscape in its resting state, a feature which is lost after ligands stabilize the ECD. A role for ECD dynamics in receptor activation and signalling has never been demonstrated. Here we have used nanobodies that reduce EGFR dynamics by simultaneous binding to domain II (DII) and III (DIII) of the ECD to investigate the role for ECD dynamics in receptor functioning. We show that binding of these nanobodies reduce the freedom of movement of the hinge region between DII and DIII. Concomitantly, we observe a decrease in affinity of EGF binding, inhibition of kinase activation and a block of receptor internalization. By introducing point mutations in the nanobodies, we are able to increase the dynamic state of the ECD. Now we observe again low affinity binding, but a restoration of kinase activation, and enhancement of receptor internalization. Our findings demonstrate the requirement for appropriate ECD dynamics for EGFR functioning.

INTRODUCTION

The EGFR receptor is present at the cell surface of resting cells in a dynamic state where it samples through the conformational landscape[1]. This enables the formation of inactive pre-dimers at the cell surface, which can occur with different conformational settings of the ECD described as stalk-to-stalk, head-to-head[2,3]. EGF binds with high affinity to both domain I (DI) and DIII of the ECD simultaneously, which reduces ECD dynamics leading to a stabilization of the open conformation and back-to-back receptor dimerization[4,5]. In this conformation of the EGFR dimer, the two binding sites for EGF form together the high affinity EGF binding cavity. EGFR dimers trans- or crossphosphorylate tyrosine residues on the C-terminal tails resulting in binding of signal transducer proteins therebye activating intracellular signalling cascades and receptor internalization. These intracellular signalling pathways are involved in cell proliferation, migration and differentiation. Internalization is part of the down regulation process, an efficient negative feedback system leading to the degradation of EGFR in lysosomes. EGFR has frequently been found overexpressed and/or mutated in tumors explaining why this receptor has become a common target in anti-cancer therapy.

The requirement of ECD dynamics for receptor functions as ligand-binding, activation and internalization is not clear, although point mutations in the EGFR ECD of glioblastoma patients have been found to increase conformational sampling of the ECD[6]. In order to investigate the role of ECD dynamics in receptor functioning, we have used nanobodies that were previously shown to inhibit EGFR activation[7,8]. Crystal structures of these nanobodies in complex with the ECD of EGFR have revealed two types of inhibitors: EGF-competing nanobodies (7D12) and allosteric inhibitors (EgA1 and 9G8)[9]. Moreover, nanobodies were found that do not inhibit EGFR activation (EgB4)[10]. The binding epitope of 7D12 overlaps with the footprint of EGF-binding on DIII, while the binding epitopes of EgA1 and 9G8 are located on both DII and DIII at the opposite site of the EGF-binding site (Fig 1.a). The binding epitope of Egb4 has been mapped to DI[3]. Based on these data we speculated that the allosteric inhibition could be the result of a reduction of ECD dynamics, preventing the stabilization of the open conformation of EGFR[9]. In order to test this hypothesis, we here used solid-state NMR (ssNMR) spectroscopy and cell biological methods to investigate the effect of allosteric nanobodies on both ECD dynamics and receptor functioning.

The nanobodies used in this study (7D12, EgB4, EgA1 and 9G8) were purified and tested for their effect on EGF binding and inhibition of receptor activation (Supplementary Figure 1.a-f). These experiments confirm that binding of EGF, complexed to Alexa Fluor 647 (EGF-A647), is not affected by EgB4 (Fig 1.b), whereas increasing concentrations of 7D12 prevent EGF-A647 binding completely (Fig 1.c). The allosteric nanobodies EgA1 and 9G8 did not completely block EGF-A647 binding (Fig 1.d, e). Pre-incubation of A431 cells with saturating concentrations of EgA1 or 9G8 resulted in a reduction of binding affinity for EGF-A647 (Fig 1.f, Supplementary Figure 1.e). As a second read-out for EGFR functioning, we tested the effect of the allosteric nanobodies on receptor internalization. Both EgA1 and 9G8 were able to block EGFR internalization as long as they remain bound to EGFR (Fig 1.g-i). Taken together, the allosteric nanobodies Ega1 and 9G8 nanobodies reduce EGF binding affinity and block both receptor activation and receptor-mediated endocytosis.

To directly test the effect of the allosteric nanobodies on ECD dynamics we used NMR spectroscopy (ssNMR) as described previously[1]. A431 cells were cultured on growth medium containing [^{13}C,^{15}N]-isotope labelled amino acids to enrich proteins for NMR-active

isotopes. The membrane fraction was isolated resulting in a sample with the EGFR as the most abundant protein[1]. To probe ECD dynamics we recorded ssNMR spectra using cross polarization (CP) and double ^{13}C quantum single quantum (DQSQ) correlation experiments that in the initial rate regime are sensitive to receptor mobility. As a reference we conducted direct excitation experiments that report on the overall amount of sample in the NMR setup (see Materials & Methods).

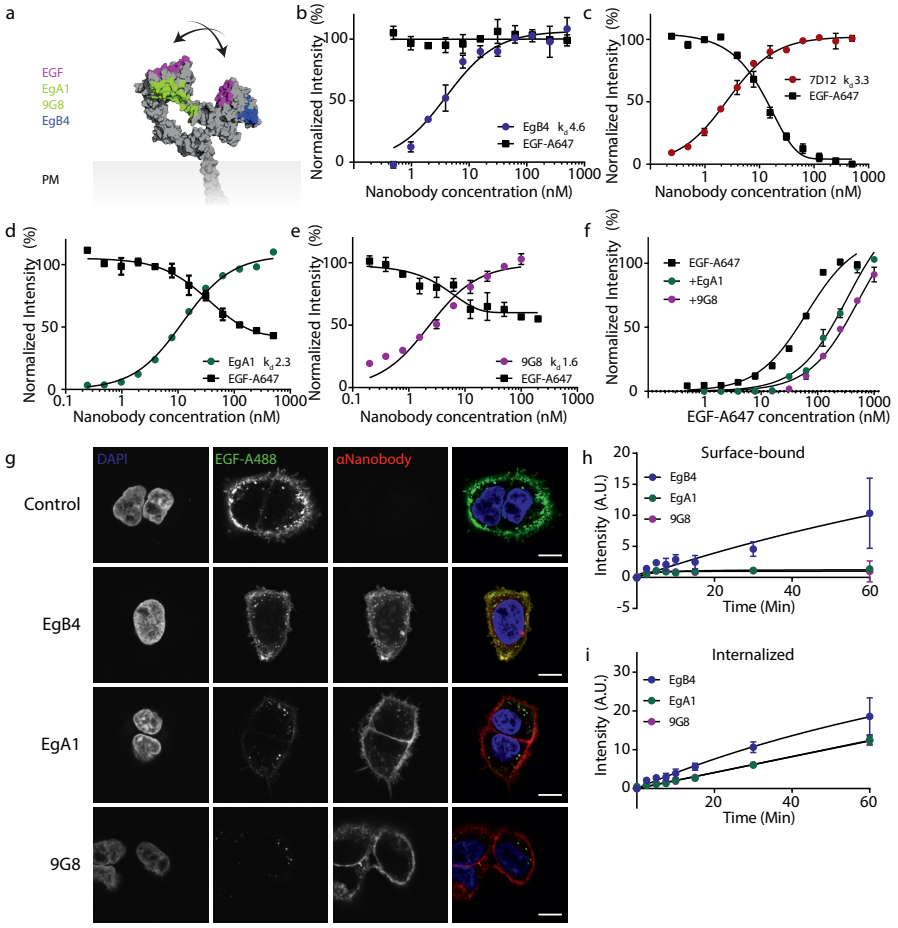


Figure 1 | Allosteric nanobodies block EGFR functioning. a, binding epitopes of αEGFR ligands, EGF (magenta; PDB accession code 1IVO), EgB4[3] (blue), 7D12 (Magenta DIII; PDB accession code 4KRL), EgA1 (green; PDB accession code 4KRO) and 9G8 (green; PDB accession code 4KRP), highlighted on a dynamic EGFR ECD (gray; PDB accession code 4KRO). **b**, Non-inhibiting nanobody EgB4 does not prevent EGF-A647 binding, **c**, while competitive inhibitor 7D12 competes for EGF-A647 binding. **d**, **e**, Allosteric nanobodies EgA1 (**d**) and 9G8 (**e**) reduce EGF-A647 binding partially. **f**, Binding affinity of EGF-A647 is reduced after binding of allosteric nanobodies EgA1 or 9G8. **g-i**, EGFR internalization assays reveal that as long as allosteric nanobodies are bound to EGFR, EGF binding does not result in receptor internalization.

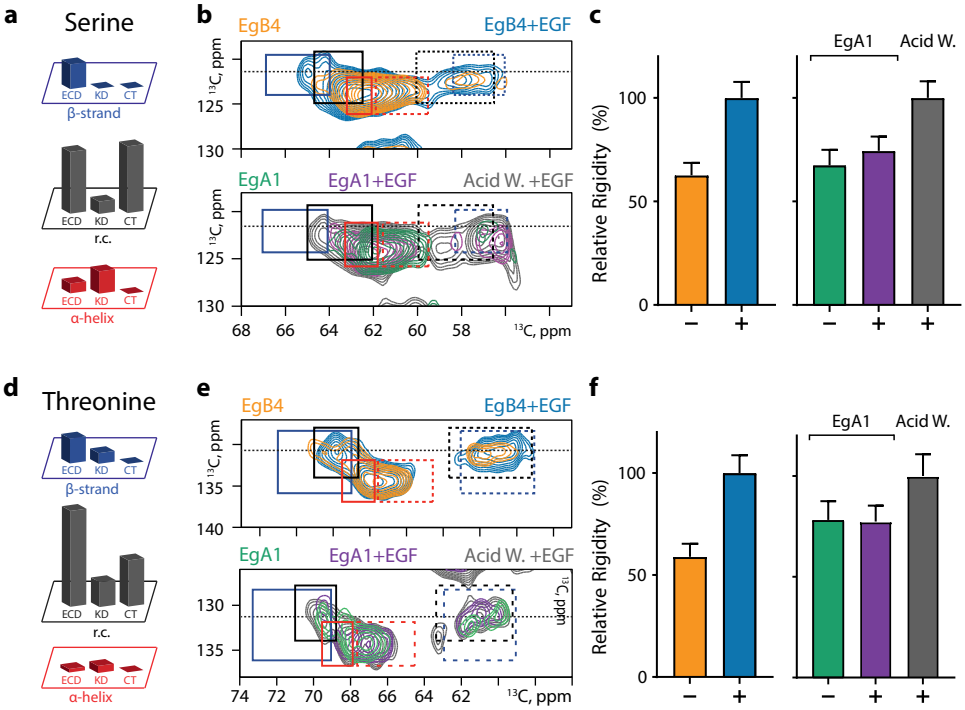


Figure 2 | Allosteric nanobody EgA1 locks DII/DIII hinge, thereby reducing ECD dynamics. a, distribution of secondary structure elements of serine residues per dubdomain of EGFR (ECD: extracellular domain; KD: kinase domain; CT: C-terminal tail). **b**, up, serine region of two-dimensional DQSQ spectrum, recorded on EgB4-bound plasma membrane vesicles, before (orange) and after (blue) addition of EGF or, **b**, down, 2D DQSQ spectrum recorded on EgA1-bound plasma membrane vesicles before (green) or after (purple) addition of EGF. An acid wash was performed to remove ligands from EGFR, followed by addition of EGF (gray). Coloured boxes indicate confidence intervals for certain secondary structure elements, e.g. α-helices (red), random coil (black) and β-sheet (blue). Dotted boxes represent Cα correlations, while solid boxes represent Cβ correlations. ssNMR spectra were recorded at a set temperature of 283 K and 9 kHz MAS. **c**, Quantification of the β-sheet peaks in a 1D slice (dashed line) which are normalized to the activated state for, left, EgB4-bound- and, right, EgA1-bound vesicles. **d-f**, presents the same analysis and data as **a-c**, but for threonine residues (**d**), NMR spectral region (**e**) and peak intensity quantification (**f**).

Incubation of A431 plasma membrane vesicles with a saturating concentration (500 nM) of control nanobody EgB4 prior to ssNMR analysis did not affect EGFR dynamics as compared to control vesicles (Supplementary Figure 2.a). As shown previously, addition of EGF induces an increase in rigidity in the sample as seen by increased sensitivity for rigidity-sensitive CP and DQSQ and a loss of signal in motion-sensitive INEPT spectra (Supplementary Figure 2.b-d). For improved analysis of the effect of EGF on ECD stabilization, 2D $^{13}C,^{13}C$-correlation DQSQ spectra were recorded, which provide information on changes in rigidity for secondary structure elements on a ms time scale (Supplementary Figure 3.a). Four specific regions were used for data interpretation as these regions were

resolved and comprise resonances from serine-, threonine-, alanine- and proline residues (Fig 2; Supplementary Figure 3.a-e). EGF is able to induce a stabilization of the ECD on EgB4-bound A431 plasma membrane vesicles as seen by an increased signal for β-sheet (blue boxes) and random coil (black boxes) secondary structure elements, further highlighted by 1D slices and quantification of peak intensity (Fig 2, Supplementary Figure 4). Taken together, these data show the absence of any significant effect of EgB4 binding on EGFR dynamics. The DI-specific EgB4 nanobody will be used as control throughout the following studies.

To investigate the impact of EgA1 on ECD dynamics, we incubated [^{13}C,^{15}N]-isotope-labelled A431 plasma membrane vesicles with 500 nM EgA1 for 60 min and analyzed the vesicles using ssNMR. The addition of EgA1 does not have a clear impact on the overall dynamic state of the ECD (Supplementary Figure 5.a-c). However, two-dimensional DQSQ spectra revealed that ECD-specific peak intensities are increased as compared to inactivated receptors, albeit reduced compared to fully stabilized ECDs (Fig 2; Supplementary Figure 6.a-e). These data suggest that the ECD still exhibits motion, even though the rotational freedom around the hinge region between DII and DIII is locked. Addition of EGF to the EgA1-bound plasma membrane vesicles did not induce a significant difference in the dynamic state of the sample (Supplementary Figure 5.a-c). Interestingly, also in the 2D DQSQ spectra no additional stabilization of the ECD was observed after addition of EGF, suggesting that, even though EGF is able to bind EGFR, it cannot stabilize the ECD further (Fig 2). To confirm that this is caused by EgA1, we performed an acid wash on the plasma membrane vesicles to remove all bound ligands from the sample, which shows again the EGF-induced ECD stabilization after EGF addition (Fig 2; Supplementary Figure 6). In conclusion, EgA1 locks the DII/DIII-hinge region which reduces the conformational freedom of the ECD.

EgA1 binds both the DIII and DII of the ECD, of which the latter interaction is mediated by a positively-charged arginine and hydrophobic tyrosine at position 107 and 108 respectively[9] (Supplementary Figure 7.a). By creating two EgA1 mutants, EgA1-R107D (EgA1-D_1) and EgA1-R107D,Y108D (EgA1-D_2), we aimed at abolishing the interaction of EgA1 with DII by introducing negative charges on the nanobody CDR3 loop, without disrupting the DIII binding site (Supplementary figure 7.b, c). Binding assays of the EgA1 mutants confirm the expected reduction in affinity of the two mutations while they still compete for binding with wild type EgA1 (Supplementary figure 7.d-g). We performed ssNMR experiments with EgA1-D_2-bound plasma membrane vesicles to test the effect of this mutant on ECD dynamics. Interestingly, binding of EgA1-D_2 potentially increased the conformational dynamics of the ECD prior to EGF binding, which is subsequently fully stabilized by the binding of EGF (Fig 3.a-d, Supplementary Fig 7). EGF-induced tyrosine activation as well as the internalization rate constant were increased by EgA1-D_2 (Fig 3.e,f). However, increased ECD dynamics did not lead to a high affinity state of EGFR (Fig 3.g-i, Supplementary Figure 7.h-j). Taken together, the introduction of mutations in the loop of EgA1 that interacts with DII, we were able to increase the dynamic behaviour of the ECD which resulted in an increased tyrosine kinase activation and EGFR internalization, but low-affinity binding to EGF.

Our results provide novel insights into the role of the conformational freedom of the ECD for receptor functioning (Fig 4). Restricting motion around the crucial DII/DIII-hinge prevents establishment of high-affinity binding, ligand-induced tyrosine kinase activity and receptor-mediated internalization. EGFR dimerization is initiated by the ECD, and is proceeding as a closing zipper across the membrane where it brings the intracellular tyrosine kinase domains in close proximity, enabling cross-phosphorylation of the receptor.

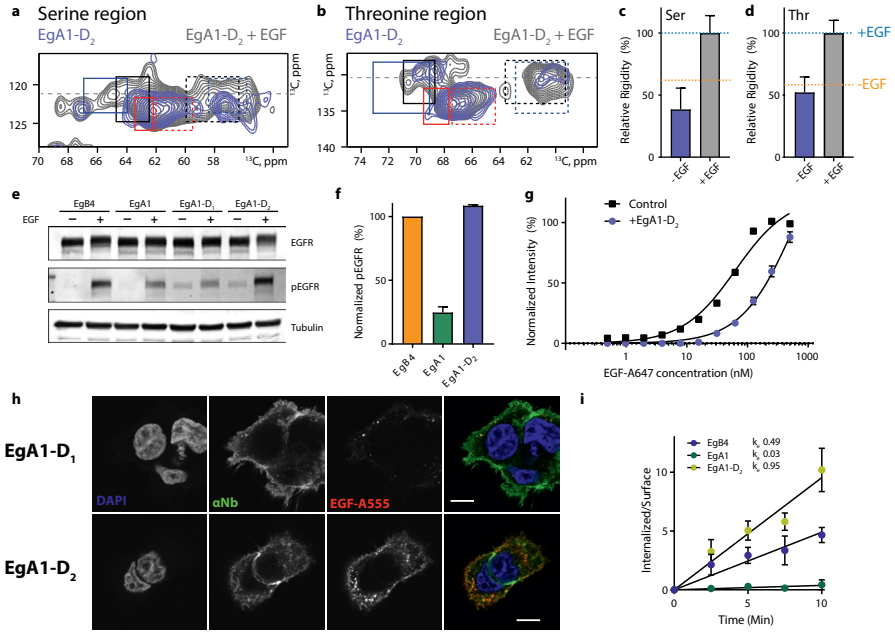


Figure 3 | ECD dynamics is crucial for EGFR functioning. a-d, Serine (**a**) and threonine (**b**) region of a two-dimensional DQSQ spectra highlighting the recovery of ECD dynamics prior to EGF binding, **c**, **d**, which is further illustrated by the quantification of β-sheet peak intensities for the serine and threonine regions respectively. Coloured boxes indicate confidence intervals for certain secondary structure elements, e.g. α-helices (red), random coil (black) and β-sheet (blue). Dotted boxes represent Cα correlations, while solid boxes represent Cβ correlations. ssNMR spectra were recorded at a set temperature of 283 K and 9 kHz MAS. **e**, **f,** Increased motion of the ECD allows for EGFR activation. g, EGF binds with lower affinity to EGFR:EgA1-D_2 complexes. **h**, **i**, EgA1-D_2 enhances EGFR internalization rates.

Subsequently, receptor dimers will oligomerize in a back-to-back fashion resulting in receptor clustering and internalization. The key step in this process is the initial formation of the ECD dimer and our data demonstrate that for this step the appropriate ECD dynamics is required. A reduction of conformational freedom results in an inhibtion of binding affinity, receptor activation and internalization. Conversely, an increase in the freedom of movement in the hinge region also discloses high-affinity binding, but stimulates tyrosine kinase activation, and leads to an enhanced receptor internalization rate. This suggests that receptors with reduced ECD dynamics do not become stabilized by EGF as a flexible hinge region is required for this process. Binding to this restrained conformation occers with a reduced affinity. Increased dynamics of the respective ECDs on the other hand is probably causing the formation of less stable, restrained dimers resulting in low-affinity binding but also in an increase of their dissociation rate after cross-phophorylation11. As a consequence, this would lead to an increase in signal propagation over the surface of the cell due to an increase of the kiss-and-run principle, previously suggested by Verveer and collaborators12. The subsequent increase in receptor phosphorylation might lead to an enhancement of signalling and receptor-mediated endocytosis.

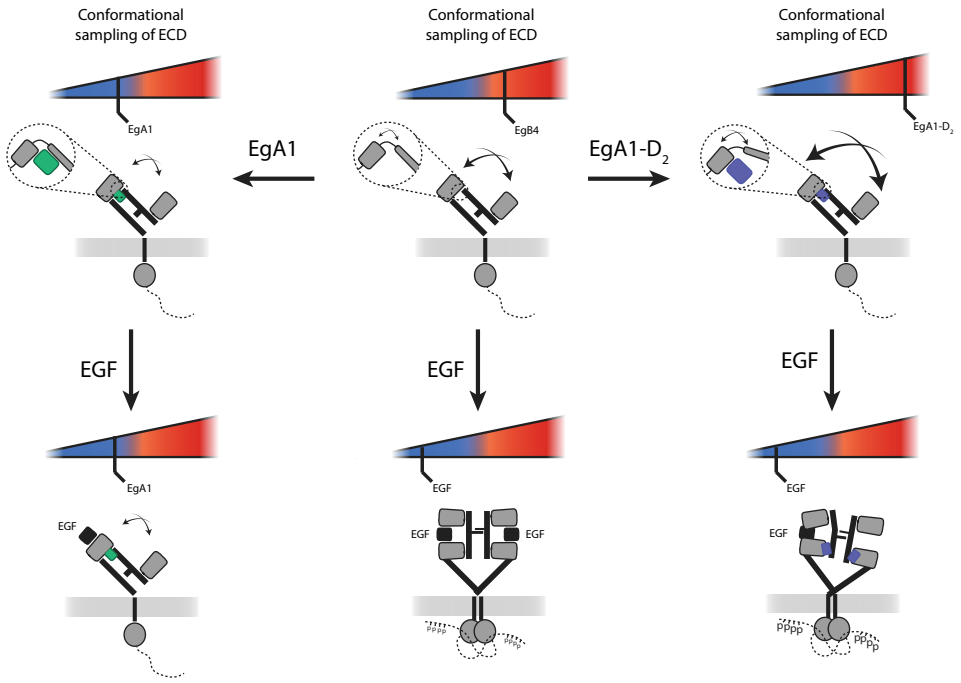


Figure 4 | Dynamics of the EGFR ECD is crucial parameter in EGFR functioning. The ECD of EGFR exhibites a high degree of motion in its resting state. The flexibility in the hinge region between DII and DIII allows for conformational exchange necessary for EGFR functioning. EGF binds to DI and DIII simultaneously, thereby stabilizing the ECD in an extended conformation leading to receptor dimerization. The dimer which is formed is symmetric. Allosteric nanobody EgA1 locks the crucial hinge region, left panel, which reduces the conformational landscape of the ECD. EGF can still bind to DIII, but it is unable to stabilize the ECD, a step which is required for EGFR functioning. The interaction site between DII of EGFR and EgA1 can be abolished with the introduction of point mutations, right panel, which result in an increased dynamic profile of the ECD. Even though the binding affinity of EGF is lower, EGFR oligomerization, activation and internalization is increased as a conseuqence of assymetric dimer formation.

Our data demonstrate that the appropriate dynamics of the ECD is crucial for proper EGFR functioning. The allosteric anti-EGFR nanobodies can be used to inhibit receptor kinase functions, but by making point mutations these nanobodies can also be used to stimulate receptor down regulation. These observations offer opportunities for the design of novel therapeutics for controlling receptor functions of the EGFR, both inhibitory as well as stimulatory. As the other members of the ErbB receptor family have the similar domain characteristics, this rationale might be extended to other receptors as well.

Materials & Methods

Cell culture. A431 cells were cultured until 80-90% confluency in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and Pen/Strep in an atmosphere containing 5% CO_2 and 37°C.

Nanobody production. *E. coli* BL21C+ cells containing plasmid DNA for his-tagged EgB4, 7D12, 9G8, EgA1, EgA1-D_1 and EgA1-D_2 were grown overnight in LB as a preculture. On the morning of the production, overnight cultures were diluted 1/500 into auto-induction medium [1/50 of 5052 (0.5% glycerol, 0.05% glucose, 0.2% α-lactose), 1/20 NPS (100 mM NaHPO4, 25 mM $(NH_4)2SO_4$, 50 mM KH_2PO_4), 2mM $MgSO_4$, micronutrients and kanamycin in LB]. The next day, the bacteria were harvested and resuspended in PBS. After several freeze-thaw cycles, the periplasm fraction was collected. The nanobodies were purified using immobilized metal-affinity chromatography (IMAC) using the His-tag. A HisTrap HP Ni^{2+}-Sepharose column was used to trap His-tag containing proteins. The nanobodies were reconsitutated at 1 mg/mL concentration for storage at 4°C.

Phosphorylation Assay. A431 cells were grown in a 6-well culture plate to 70% confluency before the medium was replaced to DMEM without FBS the day before the experiment. The cells were washed with binding buffer (DMEM–FBS buffered with 10 mM HEPES, pH 7.4) after which the cells were preincubated for 1 hour at 4°C with EgB4 (100 nM), 7D12 (100 nM), 9G8 (100 nM), EgA1 (100 nM), EgA1-D_1 (100 nM) or EgA1-D_2 (100 nM) to allow maximal binding of the nanobodies. EGF (33 nM) was added to the cells to induce EGFR phosphorylation for 1 hour at 4°C. The cells were washed twice with ice-cold phosphate-buffered saline (PBS) followed by cell lysis using a RIPA buffer (150 mM sodium chloride, 1.0% Triton X-100, 0.1% SDS, 50 mM Tris pH 8.0 and a protease inhibitor cocktail (Complete, Roche)). Proteins were separated by SDS-PAGE followed by western blotting onto PVDF-membrane. The membrane was incubated with Mouse-α-EGFR 1F4 (Cell Signaling, Danvers, Massachusetts), Rabbit-α-phosphoEGFR Y1068 (Cell Signalling Technology, Danvers, Massachusetts) and Mouse-α-Tubulin overnight at 4°C. The antibody staining was visualized using Goat-α-Rabbit IRDye700 (Li-COR) and Goat-α-Mouse IRDye800 (Li-COR) and detected on an Odyssey imaging system (Li-COR).

Binding Assay. A431 cells were grown in a 96-well Nunclon culture plate (ThermoFisher) to 80-90% confluency. The cells were washed with binding buffer (DMEM–FBS buffered with 10 mM HEPES, pH 7.4) after which the cells were incubated for 2 hours at 4°C with a serial dilution of nanobodies or EGF-A647. The cells were washed twice with ice-cold PBS before they were fixed using a 4% paraformaldehyde solution (PFA) for 15 minutes at room temperature. For nanobody detection, the cells were incubated for 1 hour at room temperature with Rabbit-α-VHH (Creative BioLabs) followed by an incubation of 1 hour at room temperature of Goat-α-Rabbit IRDye800 (Li-COR). The 96-well plate was read using an Odessey imaging system (Li-COR) and wells were quantified using the Odessey software (Li-COR) and ImageJ.

Competition Assay. A431 cells were seeded the day before the experiment at a density of 15.000 cells/well into a 96-well nunclon culture plate. The cells were incubated with a serial dilution of a nanobody of interest and allowed to bind to the cells for 1 hour at 4°C. Afterwards, saturating concentration of EGF-A647 (100 nM) or competitor nanobody (500 nM) was added to the binding solution and was again incubated for 1 hour at 4°C. The plate was directly imaged after the incubation period as fluorophores were conjugated to the ligands.

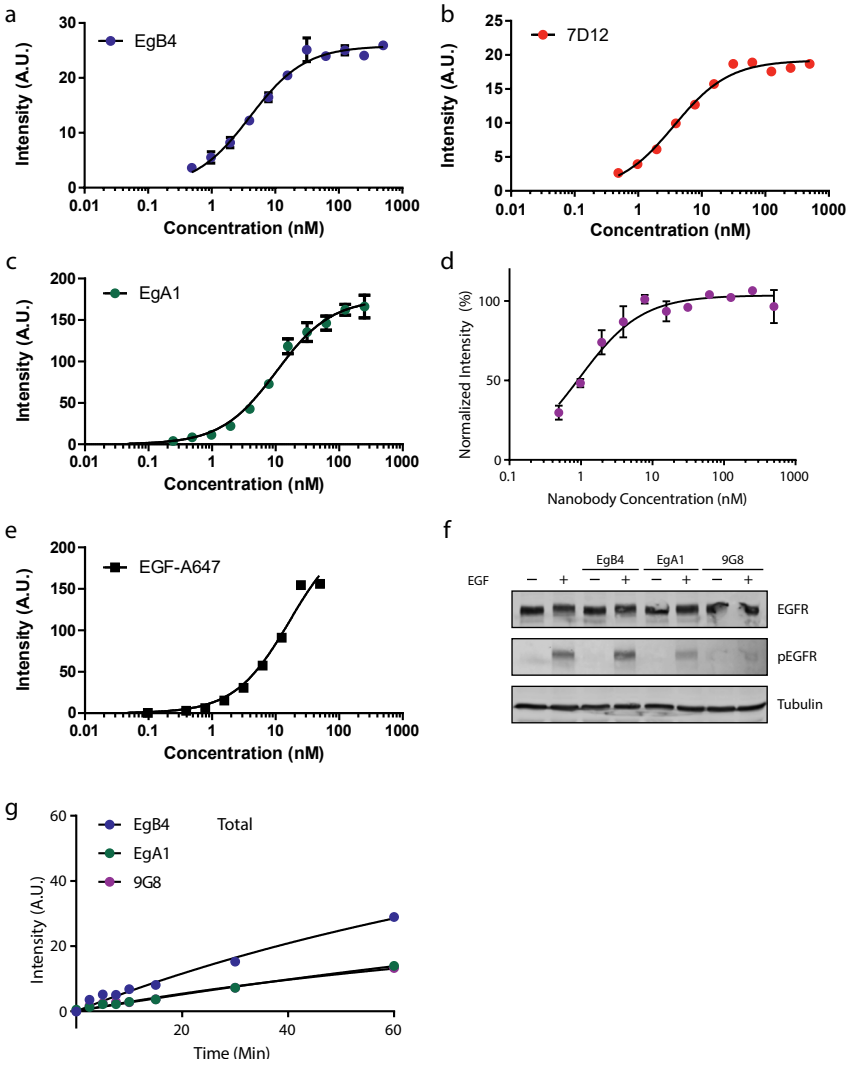
Confocal Microscopy. A431 cells were seeded the day before the experiment at a concentration of 100.000 cells/well onto 12 mm coverslips in a 12-well culture plate. The cells were preincubated with either EgB4 (100 nM) or EgA1 (100 nM) for 30 minutes at 37°C followed by the addition of EGF-A488 (100 nM, ThermoFisher) for another 10 minutes at 37°C. The cells were washed twice with ice-cold PBS and fixed using a 4% PFA solution. Confocal imaging was performed on a ZEISS LSM700 confocal microscope at 63x magnification.

Internalization assay. A431 cells were seeded prior to the experiment in a 96-well nunclon culture plate and were allowed to grow to 80-90% confluency. The cells were serum starved with DMEM without FBS 30 minutes before the experiments. Nanobodies were diluted into DMEM –FBS to a final concentration of 500 nM and were incubated on the cells for 30 min at 37°C. EGF-A647 was added for a period of 90 minutes with several intervals between each time point. After the last time point, the cells were washed twice with ice-cold PBS and were

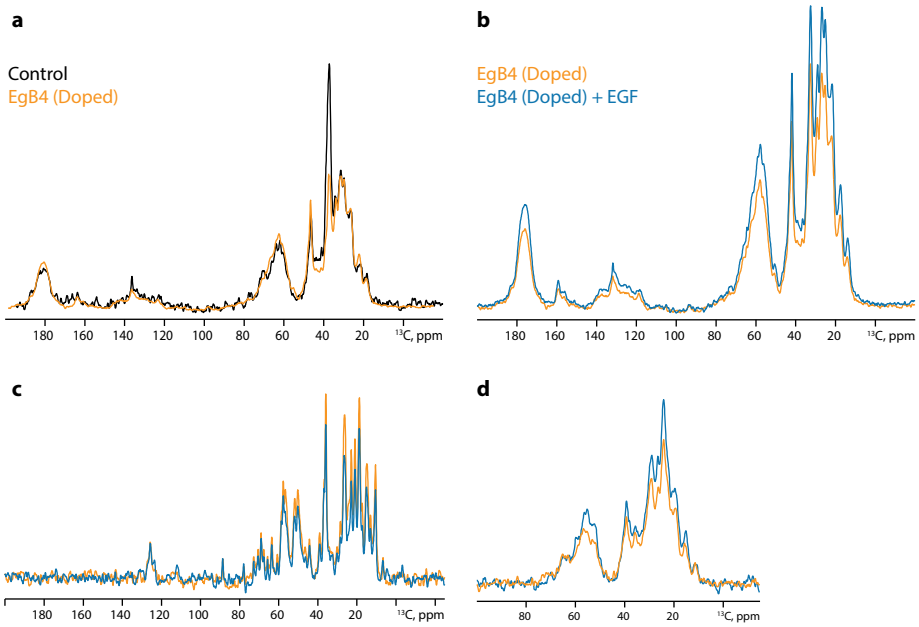
REFERENCES

1. Kaplan, M. et al. EGFR Dynamics Change during Activation in Native Membranes as Revealed by NMR. Cell (2016). doi:10.1016/j.cell.2016.10.038
2. Bessman, N. J., Bagchi, A., Ferguson, K. M. & Lemmon, M. A. Complex Relationship between Ligand Binding and Dimerization in the Epidermal Growth Factor Receptor. Cell Rep. (2014). doi:10.1016/j.celrep.2014.10.010
3. Zanetti-Domingues, L. C. et al. The architecture of EGFR's basal complexes reveals autoinhibition mechanisms in dimers and oligomers. Nat. Commun. (2018). doi:10.1038/s41467-018-06632-0
4. Ogiso, H. et al. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell (2002). doi:S0092867402009637 [pii]
5. Garrett, T. P. J. et al. Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor α. Cell (2002). doi:10.1016/S0092-8674(02)00940-6
6. Orellana, L. et al. Oncogenic mutations at the EGFR ectodomain structurally converge to remove a steric hindrance on a kinase-coupled cryptic epitope. Proc. Natl. Acad. Sci. U. S. A. (2019). doi:10.1073/pnas.1821442116
7. Roovers, R. C. et al. A biparatopic anti-EGFR nanobody efficiently inhibits solid tumour growth. Int. J. Cancer (2011). doi:10.1002/ijc.26145
8. Roovers, R. C. et al. Efficient inhibition of EGFR signalling and of tumour growth by antagonistic anti-EGFR Nanobodies. Cancer Immunol. Immunother. (2007). doi:10.1007/s00262-006-0180-4
9. Schmitz, K. R., Bagchi, A., Roovers, R. C., Van Bergen En Henegouwen, P. M. P. & Ferguson, K. M. Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. Structure (2013). doi:10.1016/j.str.2013.05.008
10. Hofman, E. G. et al. EGF induces coalescence of different lipid rafts. J. Cell Sci. (2008). doi:10.1242/jcs.028753
11. Alvarado, D., Klein, D. E. & Lemmon, M. A. Structural Basis for negative cooperativity in growth factor binding to an EGF receptor. Cell (2010). doi:10.1016/j.cell.2010.07.015
12. Verveer, P. J., Wouters, F. S., Reynolds, A. R. & Bastiaens, P. I. H. Quantitative imaging of lateral ErbB1 receptor signal propagation in the plasma membrane. Science (80-.). (2000). doi:10.1126/science.290.5496.1567

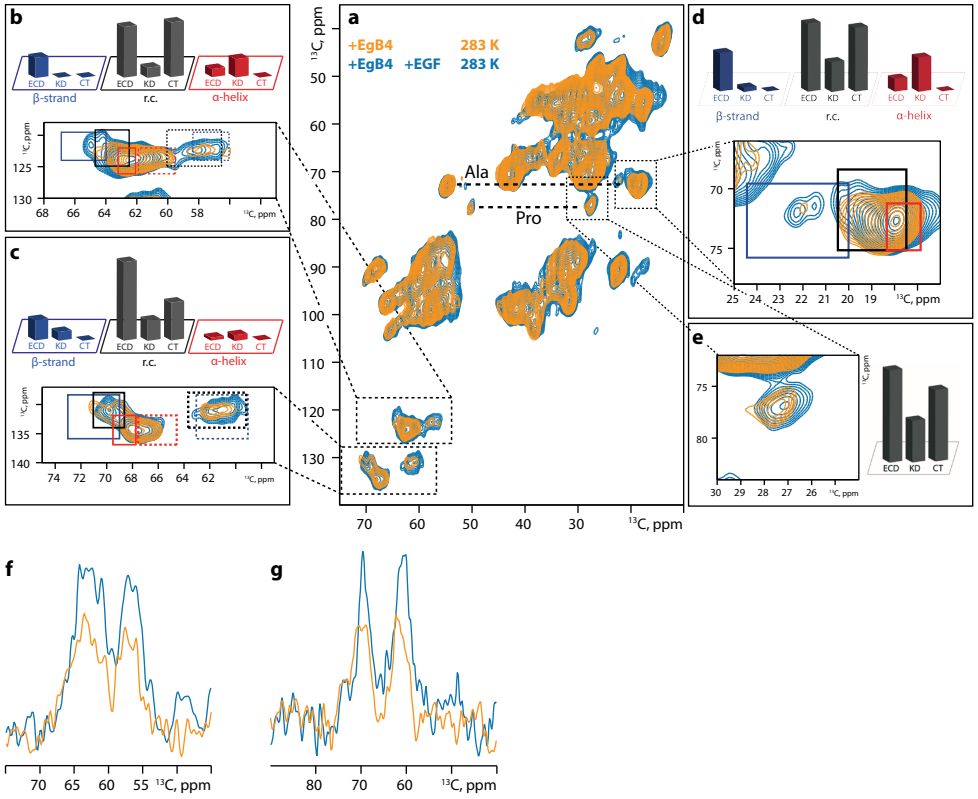
SUPPLEMENTARY INFORMATION



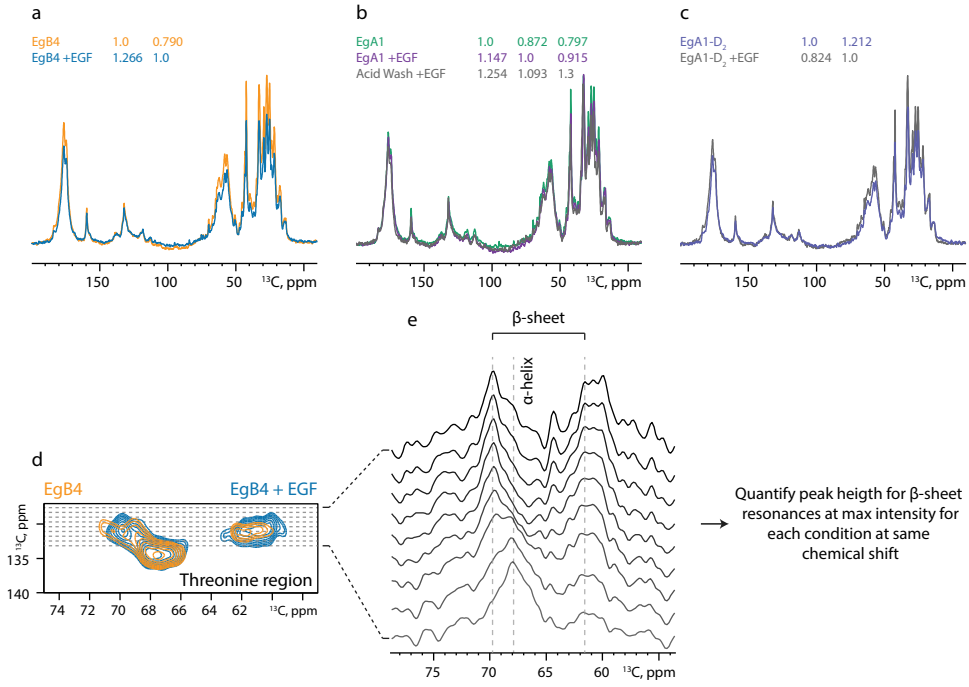
Supplementary Figure 1 | Characterization of individual EGFR ligands on EGFR. **a-e**, binding assay of individual EGFR ligands: nanobodies EgB4, 7D12, EgA1, 9G8 and EGF-A647 respectively. **f**, Allosteric nanobodies EgA1 and 9G8 (both 500 nM) reduce EGF-induced EGFR phosphorylation as compared to the absence of nanobody or control EgB4 on A431 cells. **g**, Total fluorescence intensity measured over time on A431 cells doped with EgB4 (blue), EgA1 (green) and 9G8 (magenta) after stimulation with EGF-A647.



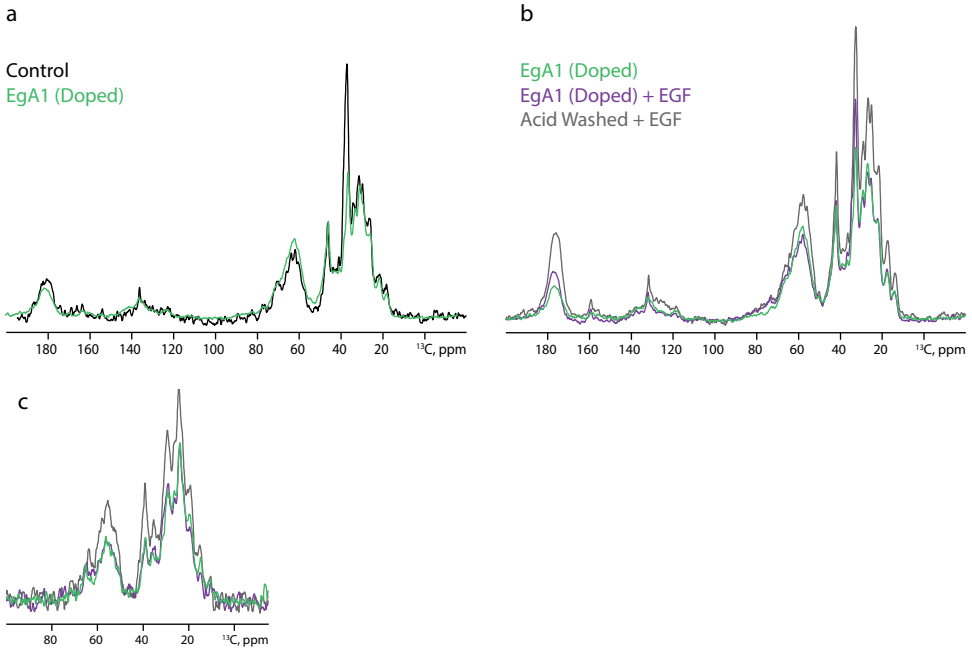
Supplementary Figure 2 | ssNMR data on EgB4-bound EGFR on A431 plasma membrane vesicles. a, Comparison of a HC CP recorded on A431 control plasma membrane vesicles (black) and vesicles after addition of EgB4 (orange). **b-d**, Comparison of a HC CP, HC INEPT and DQSQ respectively recorded on EgB4-doped A431 control plasma membrane vesicles (orange) and EgB4-doped vesicles after addition of EGF (blue).



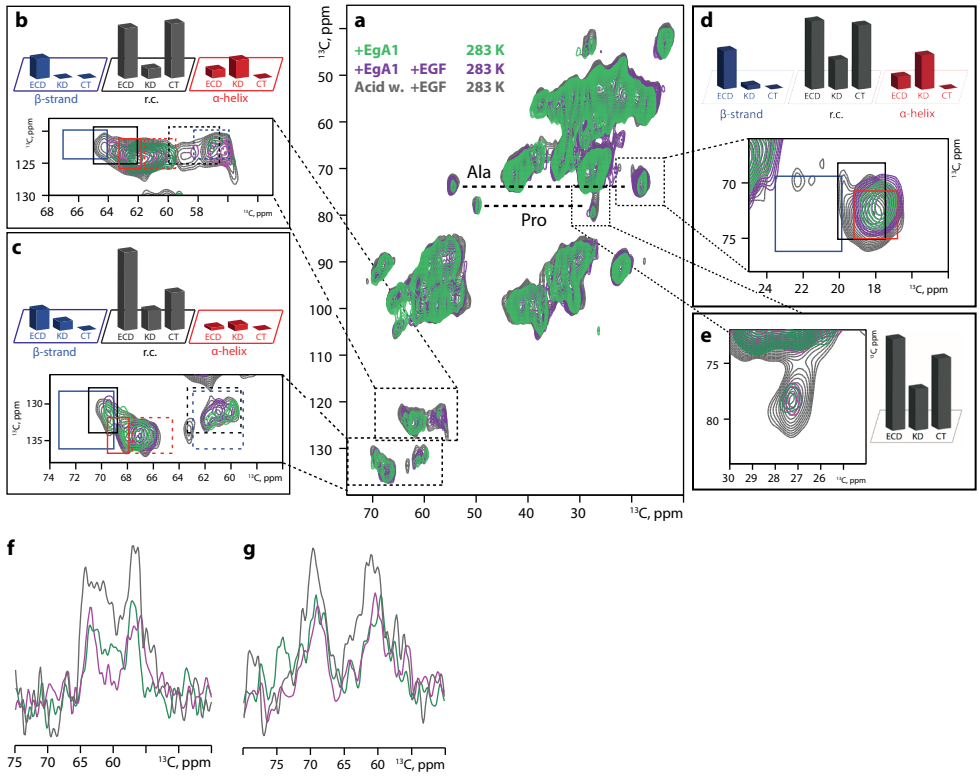
Supplementary Figure 3 | EGF stabilizes the ECD of EGFR independently of EgB4. a, DQSQ CC correlation spectra recorded on fully [^{13}C,^{15}N]-labeled EgB4-doped plasma membrane vesicles without (orange) and with (blue) EGF at ambient temperature. **b-e**, Zoom in on resolved spectral regions comprising Serine, Threonine, Alanine and Proline respectively. Coloured boxes indicated confidence intervals for α-helical (red), random coil (black) and β-sheet (blue) correlation for Cα (dashed boxes) and Cβ (filled boxes) atoms. Histograms depict the distribution of secondary structure elements per specific amino acid and domain e.g. extracellular domain (ECD), kinase domain (KD) or C-terminal tail (CT) (**b-d**) or total number of amino acids (**e**). Addition of EGF stabilizes the ECD as seen by increased sensitivity for β-sheet correlations. **f, g,** 1D slices extracted out of 2D DQSQ spectra of the serine region (**f**) and threonine region (**g**) presented in Fig 2.b, e.



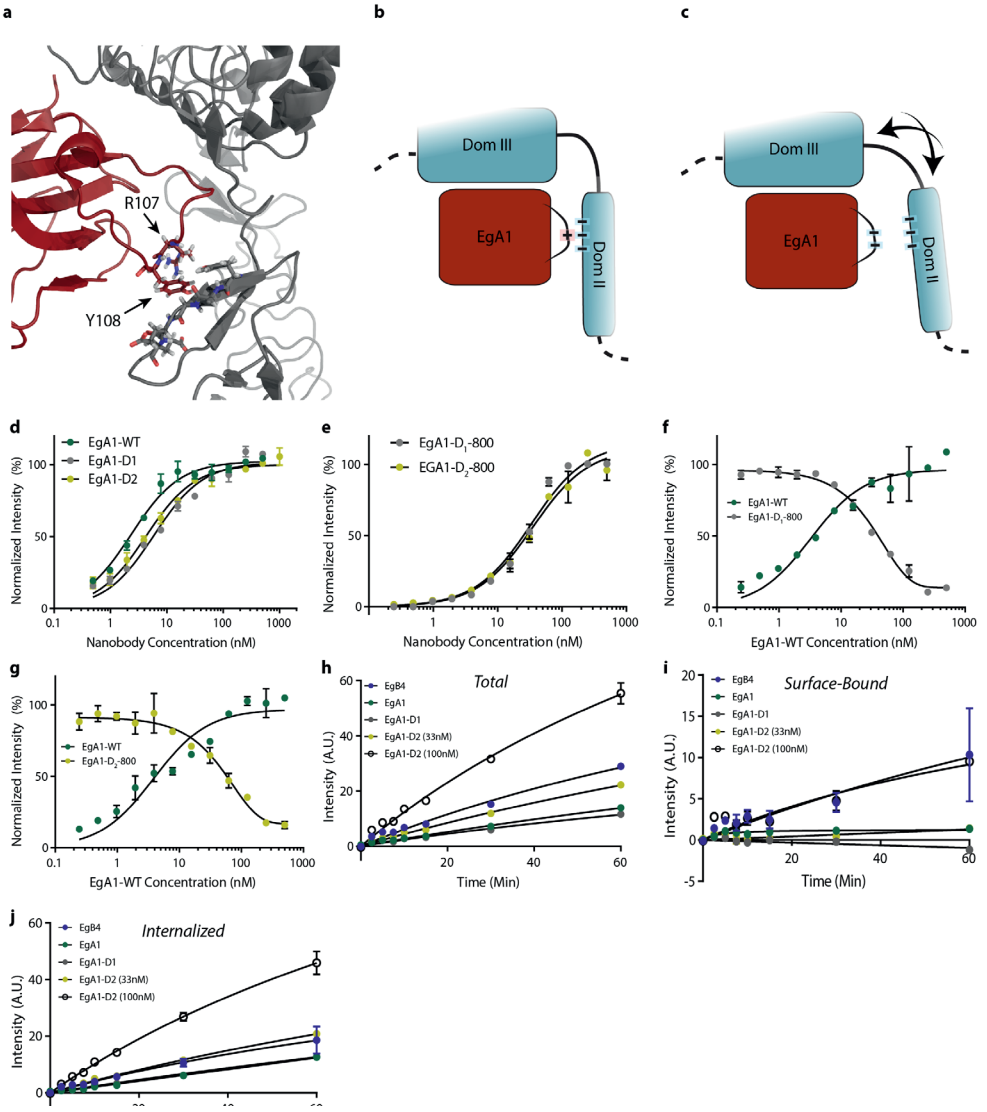
Supplementary Figure 4 | Quantification of 2D ssNMR. a-c, 1D ssNMR spectra recorded on A431 plasma membrane vesicles which are bound by **a**, EgB4 (-EGF: orange; +EGF: teal), **b**, EgA1 (-EGF: green; +EGF: purple; Acid Wash +EGF: gray) or **c**, EgA1-D2 (-EGF: blue; +EGF: gray). Spectra recorded are insensitive for motion and are used to correct for differences in sample quantity. Scaling factors between samples within the same batch are presented on top. **d**, 2D DQSQ CC spectrum of the threonine region of EgB4-bound A431 plasma membrane vesicles. 1D slices of the 2D spectra are shown, **e**, as an example for the quantification of β-sheet peak intensities.



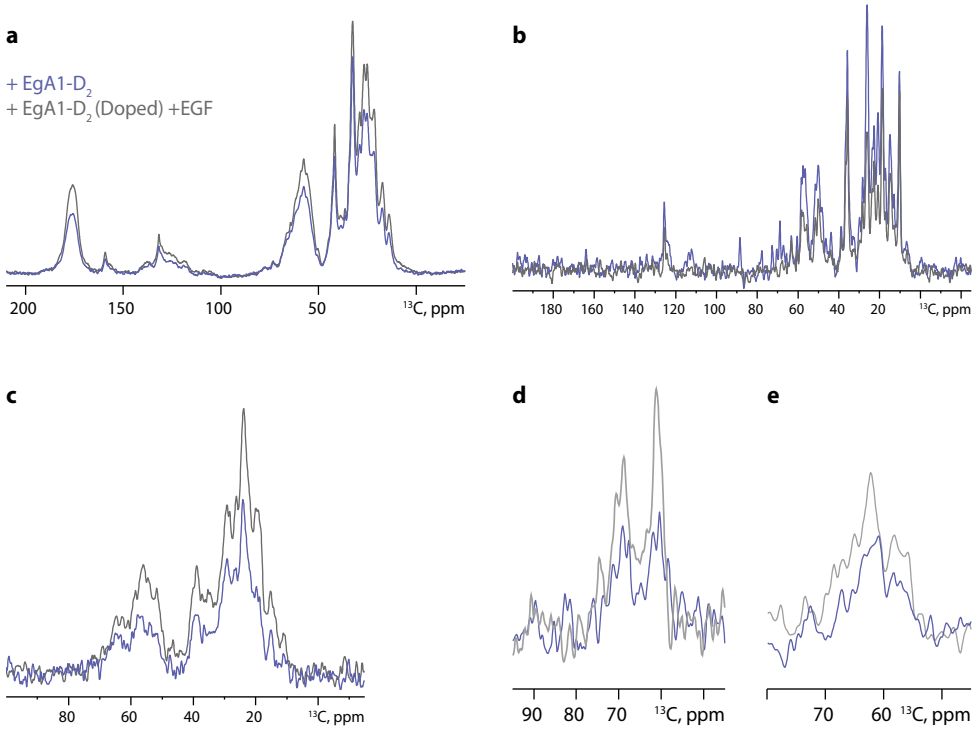
Supplementary Figure 5 | ssNMR data on EgA1-bound EGFR on A431 plasma membrane vesicles. a, Comparison of a HC CP recorded on A431 control plasma membrane vesicles (black) and vesicles after addition of EgA1 (green). **b**, **c**, Comparison of a HC CP and DQSQ respectively recorded on EgA1-doped A431 control plasma membrane vesicles (green) and EgA1-doped vesicles after addition of EGF (purple). After an acid wash, EGF was added again to the sample (gray).



Supplementary Figure 6 | EgA1 reduces the conformational freedom of the ECD of EGFR preventing EGFR activation. a, DQSQ CC correlation spectra recorded on fully [^{13}C,^{15}N]-labeled EgA1-doped plasma membrane vesicles without (green) and with (purple) EGF at ambient temperature. An acid wash was performed to remove any bound EgA1 and EGF and was reincubated with EGF (gray). **b-e,** Zoom in on resolved spectral regions comprising Serine, Threonine, Alanine and Proline respectively. Coloured boxes indicated confidence intervals for α-helical (red), random coil (black) and β-sheet (blue) correlation for Cα (dashed boxes) and Cβ (filled boxes) atoms. Histograms depict the distribution of secondary structure elements per specific amino acid and domain e.g. extracellular domain (ECD), kinase domain (KD) or C-terminal tail (CT) (**b-d**) or total number of amino acids (**e**). **f**, **g**, 1D slices extracted out of 2D DQSQ spectra of the serine region (**f**) and threonine region (**g**) presented in Fig 2.b, e.



Supplementary Figure 7 | Point mutations in CDR3 of EgA1 enhance EGF-induced EGFR activation. **a**, Zoom in on the binding site of EgA1 (red) with domain II (right) and domain III (top) of the ECD of EGFR (gray, PDB accession code 4KRO). Highlighted are R108 and Y109 of EgA1 which closely interact with a glutamate-rich sequence in domain II of EGFR. **b**, schematic representation of the ionic lock of EgA1 and the hinge region between domain II and domain III **c**, highlighting the mutation site which allows for increased mobility around the hinge region. **d-g**, EgA1 point mutants (R107D is EgA1-D_1; R107D,Y108D is EgA1-D_2) exhibit a slightly reduced affinity compared to wild type EgA1 and compete for binding to EGFR on A431 cells, confirming overlapping binding epitopes. **h**, Total fluorescence, **i**, surface-bound, and, **j**, internalized intensity measured over time on A431 cells doped with EgB4 (blue), EgA1 (green), EgA1-D_1 (gray), EgA1-D_2 (yellow, [EGF-A647]=33nM) and EgA1-D_2 (open circles, [EGF-A647]=100nM).



Supplementary Figure 8 | ssNMR data on EgA1-D_2-bound EGFR on A431 plasma membrane vesicles. **a-c**, Comparison of a HC CP, HC INEPT and DQSQ respectively recorded on EgA1-D_2-doped A431 plasma membrane vesicles (light blue) and EgA1-D_2-doped vesicles after addition of EGF (gray). **d**, **e**, Slices of the β-sheet region of the DQSQ (Fig 3.a, b) for threonine (**d**) and serine (**e**).



Chapter 4

Rigidification of the extracellular domain of EGFR stabilizes HER2 and stimulates its functioning

Reinier Damman[1,2], Apfrida Kendek[1,2], Sander van Doorn[3], Simone Lemeer[3], Paul van Bergen en Henegouwen[2], and Marc Baldus[1]

[1] NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.
[2] Cell Biology, Neurobiology and Biophysics, Department of Biology, Faculty of Science, Utrecht University, Utrecht, the Netherlands.
[3] Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.

ABSTRACT

HER2 is a member of the ErbB receptor tyrosine kinase (RTK) family and is involved in cell proliferation, differentiation, migration and inhibition of apoptosis. Upregulation of HER2 is implicated in tumorigenesis, particularly in breast cancer, and is therefore a common target in anti-breast cancer therapy. As HER2 lacks a ligand, HER2 activation is activated by heterodimerization with other members of ErbB family of RTKs, such as EGFR, HER3 and HER4. Yet, the details of how heterodimerization leads to activation of the HER2 kinase are currently elusive. Here we have investigated the effect of EGFR activation on the dynamics of the HER2 ectodomain using solid-state NMR spectroscopy. Our results suggest that the extracellular domain (ECD) of nonstimulated HER2 exhibits sizeable protein dynamics in native plasma membrane vesicles obtained from SKBR3 cells. EGF binding of EGFR leads to rigidification of the EGFR ECD, which in its turn stabilizes the HER2 ECD. Additionally, partial stabilization of the EGFR ECD by binding of a non-activating nanobody EgA1 was found to fully stabilize the HER2 ECD as seen by ssNMR and induce HER2 phosphorylation. Together, these data show that rigidification of the EGFR extracellular domain drives a similar reduction in HER2 ECD dynamics and phosphorylation of the intracellular domain.

INTRODUCTION

The Human Epidermal growth factor Receptor 2 (Her2) is a member of the receptor tyrosine kinases (RTKs) together with EGFR (HER1), HER3 and HER4[1]. Stimulation of these receptors leads to activation of intracellular signalling pathways which drive proliferation, differentiation, migration and inhibition of apoptosis. These processes are crucial for cellular development, but are required to be tightly regulated as dysregulation of HER2 is often associated with tumorigenesis, in particular in breast cancer[2,3]. Overexpression of HER2 enhances activated complexes contributing to prolonged signalling events[4–6]. Moreover, due to the constitutively active kinase domain (KD) of HER2, dimerization is sufficient to induce activation[7]. Therefore, HER2 is a common target in breast cancer therapy[8].

The overall architecture of the EGFR family is well conserved, containing an extracellular ligand-binding domain (ECD), a single transmembrane-spanning helix, a KD and a long unstructured C-terminal tail[9–13] (Fig 1.a). The ECD is further divided into 4 subdomains, DI-DIV, of which DI and DIII are involved in ligand binding, while DII and DIV are important for receptor dimerization. Interestingly, even though the global structure is conserved between these receptors, no activating ligands for HER2 have been identified yet, likely due to an additional loop in DI that occupies the ligand-binding cavity between DI and DIII. The activation of HER2 therefore relies on heterodimerization with mainly other, ligand-activated, EGFR family members. Cross-phosphorylation of the C-terminal tails of HER2 by the KD will subsequently initiate intracellular signalling cascades.

Binding of EGF to DI and DIII simultaneously to EGFR will stabilize an extended "open" conformation of the ECD which exposes dimerization motifs on DII and DIV that drive dimerization[14]. The crystal structure of the soluble ECD of HER2 suggests that it adapts an open conformation in the absence of interaction partners which would allow HER2 to readily heterodimerize. This hypothesis is further supported by molecular dynamics simulations of HER2 in bilayers which show a similar extended conformation without sizable motion of the HER2 ECD[15].

Even though several structural models have been proposed based on available crystal structures of the EGFR receptor family, the HER2 heterodimerization mechanism remains poorly understood. When comparing the structural elements of DII in the extended conformation, DII of HER2 resembles closely the DII of *D. melanogaster* EGFR in low-affinity complex to Spitz, its EGF analogue[16,17]. It is thought that the EGFR:HER2 heterodimer is therefore asymmetric and potentially unstable. In addition, these models are based on truncated constructs which are studied isolated from its native environment which may have a significant impact on HER2 functioning. It has been shown that the ECD of full-length EGFR in native plasma membranes samples through many conformations prior to ligand binding, which is crucial for EGFR functioning[18] (see chapter 3 and Damman et al. in preparation). The essential role of ECD dynamics in EGFR functioning is further underscored by the structural heterogeneity observed for EGFR dimers after ECD stabilization by high- and low-affinity ligands[19].

Previous work highlights the importance of studying full-length EGFR in its native environment for its correct functioning[20–22]. Solid-state NMR spectroscopy (ssNMR) can provide information about the structure and dynamics of membrane proteins or large protein complexes[23] and hence offers unique possibilities to study HER2 and EGFR in their native state. Recent advances in cellular- and in-cell NMR have pushed the boundaries of studying protein structure and functioning directly in cells[24–26]. In addition, ssNMR can provide insights

into changes in dynamic profile using dipolar- or scalar-based experiments[27–30]. The strength of these types of experiments was further highlighted when applied to EGFR-rich plasma membrane vesicles, which were isolated from a mammalian cell line, ensuring glycosylated functional receptors in their native lipid environment[18] (chapter 3 and Damman et al. in preparation).

We set out to adapt the plasma membrane isolation protocol for ssNMR studies for a HER2-overexpressing cell line. Our data shows that, in contrast to current models, the ECD of nonstimulated HER2 is dynamic in native plasma membrane vesicles, while the KD is rigid. After indirect activation via EGF, the HER2 ECD fully stabilizes, possibly leading to stable oligomeric complexes. Finally, stabilization of the EGFR ECD by non-activating anti-EGFR nanobody EgA1 is able to indirectly stimulate HER2 phosphorylation and induce stabilization of the HER2 ECD. Taken together, we propose a model where ligand binding to EGFR stabilizes the ECD that will function as a docking site for HER2. This view implies that rather than ligand binding, the stabilization of the ECD drives HER2 heterodimerization and the subsequent initiation of signalling events.

RESULTS

In order to study the molecular dynamics of HER2 activation in the native environment, we set out to adapt the plasma membrane isolation protocol as described elsewhere (Kaplan et al 2016). It is important that the protein of interest is one of the most abundant proteins inside the plasma membrane fraction as ssNMR measurements will be dominated by these proteins. Moreover, the relative ratio between HER2 and its activating partners should favour HER2. Therefore, we made use of the SKBR3 cell line which is known to have a high endogenous overexpression of HER2, $1.0x10^6$ HER2 receptors per cell, whilst expression levels are significantly lower levels for EGFR, HER3 and HER4 31. The relative expression levels of EGFR, HER2 and HER3 were further assessed by mass spectrometry analysis on SKBR3 cells and on the plasma membrane fraction after the isolation (Fig 1.b). In line with previous observations, HER2 is enriched after the vesiculation protocol and is present in higher numbers as compared to EGFR (~21:1), while HER3 was undetected by MS and therefore likely to be only present in low numbers. Most of the adapter proteins were lost after the vesiculation (data not shown).

Heterodimerization of HER2 with activated partners is required for the full activation of oligomeric complexes. Indirect stimulation of HER2 via EGF-activated EGFR reaches a peak after 15 minutes of continues stimulation with EGF, after which HER2 is downregulated resulting in background phosphorylation levels (Fig 1.c). SKBR3-derived plasma membrane vesicles show a reduced kinetics at which HER2 phosphorylation occurs, reaching maximal pHER2 levels after 45 minutes of EGF stimulation (Fig 1.d). Moreover, down-regulation of HER2 doesn't appear on the timescale tested. As the plasma membrane vesicles will be subjected to prolonged time periods during ssNMR experiments, it is important that stimulated HER2 complexes will remain stable during this period. EGFR-HER2 heterodimer complexes are unstable on the plasma membrane, which results in the separate internalization of EGFR and HER2. This process could be overcome by inhibiting internalization, via cytosol acidification medium, thereby maintaining EGF-bound EGFRs on the plasma membrane that continue to interact with HER2 (Fig 1.e). The latter situation resembles most likely the situation on the plasma membrane vesicles. Taken together, SKBR3-derived plasma membrane vesicles can be readily prepared for ssNMR and EGF-dependent

stimulation and phosphorylation of HER2 can be observed.

To directly assess the dynamic state of the HER2 ECD in the unstimulated- and stimulated state, we used our previous approach for protein labelling in mammalian cells which gives insights on global protein dynamics (Kaplan et al 2016). In short, SKBR3 cell were grown on modified DMEM medium, which is supplemented with [^{13}C, ^{15}N]-labelled algal amino acid mixture and 10% dialysed FBS. The cells will incorporate the labelled amino acids into their metabolism and thereby enrich proteins with NMR-active isotopes. The plasma membrane fraction was isolated and HER2 ECD dynamics were probed using a series of rigidity- (cross polarization: CP; double quantum single quantum: DQSQ) and mobility- (INEPT) sensitive ssNMR experiments. Differences in sample quantity were corrected using

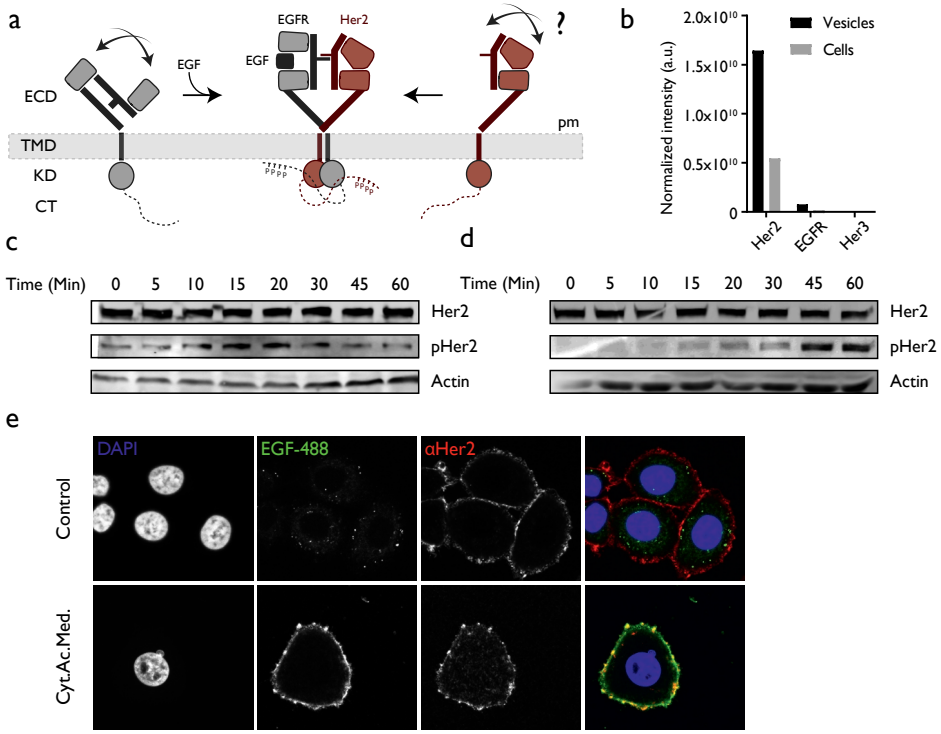


Figure 1 | HER2 forms functional complexes on SKBR3-derived plasma membrane vesicles. a, schematic representation of global structure (ECD: extracellular domain; TMD: transmembrane domain; KD: kinase domain; CT: C-terminal tail) of EGFR (gray) and HER2 (red). The EGFR ECD is dynamic prior to activation, which gets stabilized by EGF inducing heterodimerization with HER2. **b**, Mass spectrometry analysis performed on SKBR3 cells and plasma membrane vesicles derived from these cells, highlighted a enrichment for HER2 after vesiculation and HER2 being dominant over EGFR and HER3. **c**, Addition of EGF induces a peak in HER2 phosphorylation after 15-20 minutes, which is down-regulated to background HER2 phosphorylation levels. **d,** phosphorylation kinetics of HER2 are reduced as compared to cells, showing how levels of pHER2 after 45 minutes of stimulation with EGF. No downregulation of HER2 is observed. **e,** confocal imaging of nuclei (blue) EGF-488 (green) and HER2 (red) reveals that EGFR is internalized and down-regulated independently of HER2. Blocking internalization will retain EGFR-EGF-A488 complexes on the plasma membrane, allowing for continuous heterodimerization with HER2.

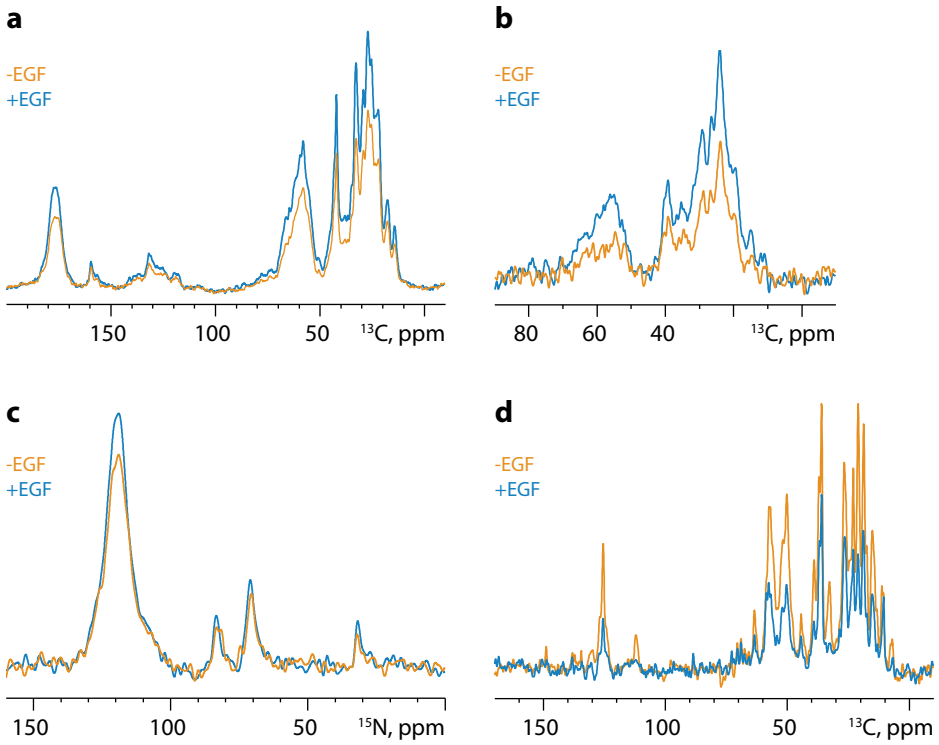


Figure 2 | EGF-stimulation induces a general reduction in motion on SKBR3 plasma membrane vesicles. a-c, addition of EGF to fully labelled SKBR3 derived plasma membrane vesicles induces an overall increase in sensitivity for rigidity-sensitive, $^{1}H^{13}C$ CP, DQSQ and $^{1}H^{15}N$ CP ssNMR spectra. d, motion-sentive $^{1}H^{13}C$ INEPT spectra show a reduction in motion after EGF addition, which is consistent with the rigidity-sensitive spectra.

90° pulses which are insensitive towards motion.

Plasma membrane vesicles were incubated for 60 minutes with EGF and show a general increase in sensitivity in rigidity-sensitive experiments as compared to unstimulated plasma membrane vesicles (Fig 2.a-c). This observation is in line with an overall decrease in signal observed for motion-sensitive experiments (Fig 2.d). These data suggest that upon addition of EGF to the plasma membrane vesicles, a general stabilization of proteins occur. To obtain a better understanding of the effect of EGF-induced rigidification, two-dimensional DQSQ experiments (as described in chapter 3) were recorded and interpreted using a statistical secondary structure analysis[18] of the available HER2 crystal structures[11,32–34] (Fig 3.a). After addition of EGF, a general increase in signal is observed for alanine and proline resonances, which are residues that are more enriched in the HER2 ECD and C-terminal tail, which suggest a stabilization of the ECD and potential oligomerization (Fig 3.b, c). Moreover, in the unstimulated state, we observe mainly α-helical correlations for serine and threonine residues, which are more predominantly found in the KD of HER2 (Fig 3.d, e). Upon addition of EGF, a general increase in signal is observed for random coil- and β-sheet secondary structure elements which indicate a stabilization of the HER2 ECD. In conclusion, these

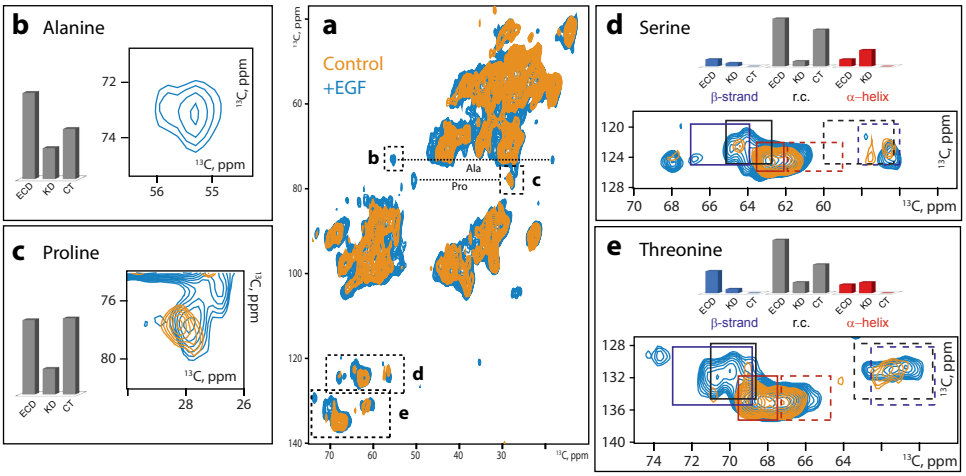


Figure 3 | HER2 ECD becomes rigid after EGF stimulation. a, two-dimensional DQSQ spectra recorded on fully labelled SKBR3 plasma membrane vesicles. **b-e**, highlights of the spectrum presented in **a**, which contain secondary structure information for alanine, proline, serine and threonine respectively. **b**, **c,** Alanine and proline residues are enriched in the ECD and C-terminal tail (CT), suggesting that the addition of EGF leads to a rigidification of the ECD and/or CT. **d**, **e**, Serine and threonine residues also display information on the secondary structure elements as highlighted by the coloured boxes. Red boxes portray confidence intervals for α-helices (Cα solid box; Cβ dashed box), black boxes for random coil structures and the blue boxes for β-sheet resonances. Stimulation with EGF induces a general increase in random coil- and β-sheet confidence intervals which are secondary structure elements predominantly found in the HER2 ECD. ssNMR spectra were recorded at 9 kHz MAS at ambient temperature.

data suggest that the ECD of HER2 exhibits sizable dynamics prior to stimulation which is in contrast with the crystal structure of the open conformation. Hence, addition of EGF leads to an EGFR-dependent stimulation and oligomerization[35] of HER2, which results in the stabilization of the ECD of HER2.

Previous work on EGFR has shown similar behaviour in receptor activation, where EGF directly stabilizes the ECD of EGFR by binding to DI and DIII (see chapter 3). To this end, no activating ligands have been identified for HER2. Hence, we postulate that the rigidification of a binding partner would be sufficient to stimulate HER2. Recently, it has been shown that allosteric nanobody EgA1 partially induces stabilization by locking a hinge between DII/DIII of the ECD[36] (and chapter 3). As long as EgA1 remains bound to EGFR, conformations that are necessary for EGF-induced stabilization do not occur, resulting in blockage of EGFR function (see chapter 3).

To ensure that αEGFR nanobody EgA1 doesn't bind to HER2 as well, we performed binding assays on EGFR-overexpressing cell line A431 and SkBr3 using EgA1 and αHER2 nanobody 11a4[37,38] (Fig 4.a). EgA1 binds with similar high affinity to EGFR on A431- (kd 7.4) and SKBR3 (kd 1.1) cells. As the affinity is similar and saturation of binding is reached at approximately 10 nM EgA1 concentration on SKBR3 cells, it is confirmed that EgA1 does not bind to HER2. Moreover, the B_{max} of 11A4 is significantly higher than the B_{max} of EgA1. Binding of control nanobody EgB4 doesn't activate HER2, where phosphorylation is only observed after EGF addition (Fig 4.b). However, the allosteric nanobody EgA1 alone induces

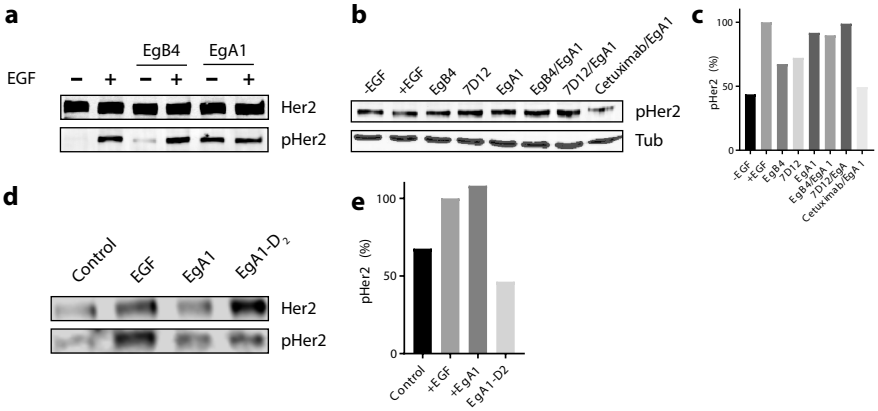


Figure 4 | Stabilization of the EGFR ECD drives HER2 phosphorylation. a, Binding assay of αEGFR nanobody EgA1 and αHER2 nanobody 11A4 performed on A431- and SKBR3 cells. **b**, phosphorylation assay highlighted the effect of nanobody binding on SKBR3 cells. The control nanobody EgB4 does not induce HER2 phosphorylation, which only occurs after EGF stimulation. EgA1 alone is sufficient to induce phosphorylation on HER2 which does not get further enhanced after addition of EGF. **c**, **d**, Preliminary phosphorylation assay, and quantification, performed on SKBR3 cells that shows that antibody Cetuximab is able to prevent EgA1-induced HER2 phosphorylation. e, f, Preliminary phosphorylation assay, and quantification, performed on SKBR3 cells which show that an EgA1 mutant (EgA1-D_2) cannot induce HER2 phosphorylation as wild type EgA1.

phosphorylation of HER2, even in the absence of EGF. These data suggest that the rigidification in the EGFR ECD after EgA1 binding forms a seed for HER2 heterodimerization. As EGF does not stabilize the EGFR ECD anymore after EgA1-binding, it is unlikely that the back-to-back dimer interface will be involved in HER2 heterodimerization. A second dimer interface, front-to-front dimer, is described in a tetramer model, where the 2 central EGFR molecules are involved in the oligomerization but do not have any ligand bound to their respective ECDs. In this model, DI and DIII are involved in the interactions that give rise to the tetramer[39]. Preliminary results reveal that DIII of EGFR is involved in the heterodimerization process, but in a size-dependent manner, as nanobody 7D12 is not able to block HER2 activation. However Cetuximab, a monoclonal antibody which has an overlapping binding epitope with 7D12, does block HER2 activation (Fig 4.c, d). To confirm that HER2 activation is driven by stabilization of the EGFR ECD, we introduced point mutations into EgA1 (R107D, Y108D; EgA1-D_2). Due to these mutations, we lose interactions of CDR3 of EGA1 with DII of EGFR, thereby freeing the hinge between DII/DIII and subsequently rescuing ECD dynamics (see chapter 3). Interestingly, our preliminary results show that EgA1-D_2 does not activate HER2 anymore (Fig 4.e, f), suggesting that the stabilization of the EGFR ECD drives HER2 activation.

To test the effect of EgA1 on HER2 stabilization we set out to record ssNMR spectra on SKBR3 plasma membrane vesicles. The addition of EgA1 induces a general stabilization as observed in rigidity-sensitive spectra (Fig 5.a-c) and a loss of signal in motion-sensitive spectra (Fig 5.d). These data are in line with the ssNMR data where EGF was used to activate HER2 (Fig 3) and the biochemical data on HER2 activation (Fig 4). Two-dimensional DQSQ

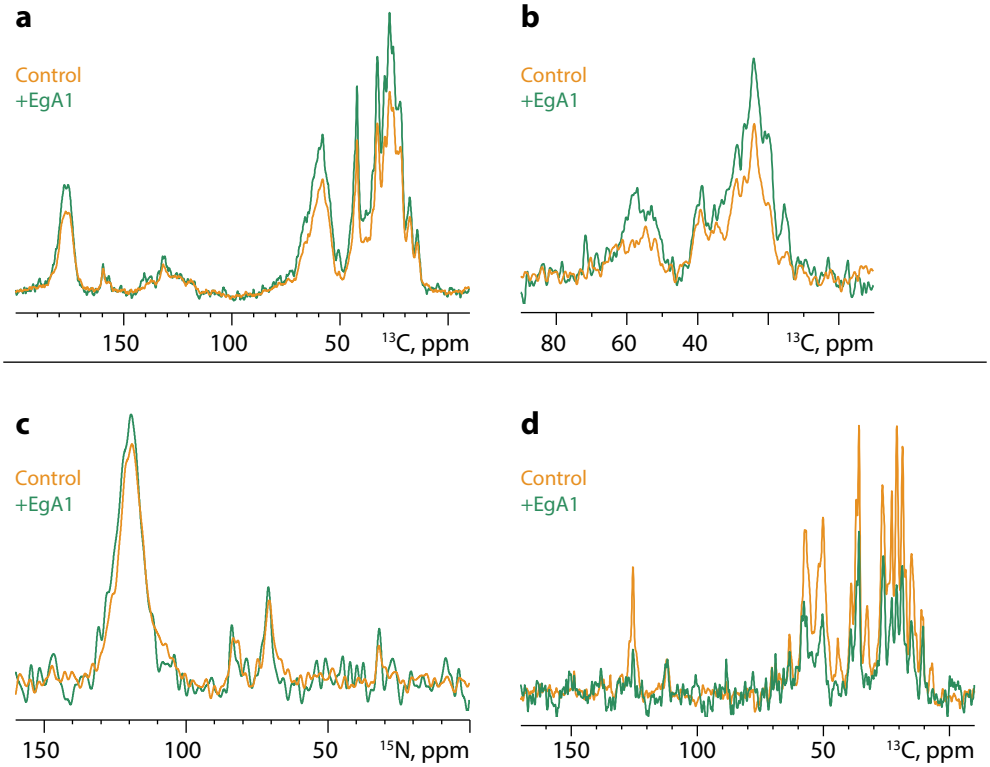


Figure 5 | EgA1 induces an overall rigidification on SKBR3 plasma membrane vesicles. a-c, addition of EgA1 to fully labelled SKBR3 derived plasma membrane vesicles induces an overall increase in sensitivity for rigidity-sensitive, ^{1}H^{13}C CP, DQSQ and ^{1}H^{15}N CP ssNMR spectra. **d**, motion-sentive ^{1}H^{13}C INEPT spectra show a reduction in motion after EgA1 addition, which is consistent with the rigidity-sensitive spectra.

spectra (Fig 6.a) reveal a strong increase in signal for both proline, serine and threonine region by EgA1 (Fig 6.b-d). Using the statistical analysis as described above, we observed increased signals for random coil and β-sheet resonances for proline, serine and threonine residues (Fig 6.b-g). Taken together, these results would be consistent with the notion that EgA1 partially stabilizes the ECD of EGFR, where it will function as a docking site for HER2 to bind to and initiate ECD stabilization, oligomerization and transphosphorylation.

DISCUSSION

Previous evidence has shown that studying the molecular activation of growth factor receptors in native plasma membranes is desirable in order to derive accurate models on their dynamical properties and functioning (Kaplan et al., 2016, Damman et al. in preparation). The native environment imposes external factors which are lost under *in vitro* conditions, but can have major implications on receptor functioning. ssNMR has proven to be powerful as motion and rigidity can be probed in succession on complex fully [^{13}C, ^{15}N]-labelled plasma membrane vesicles and large protein structures, moving closer toward (near) in-cell settings. Here, we adapted an existing protocol for protein labelling in mammalian cell lines for a

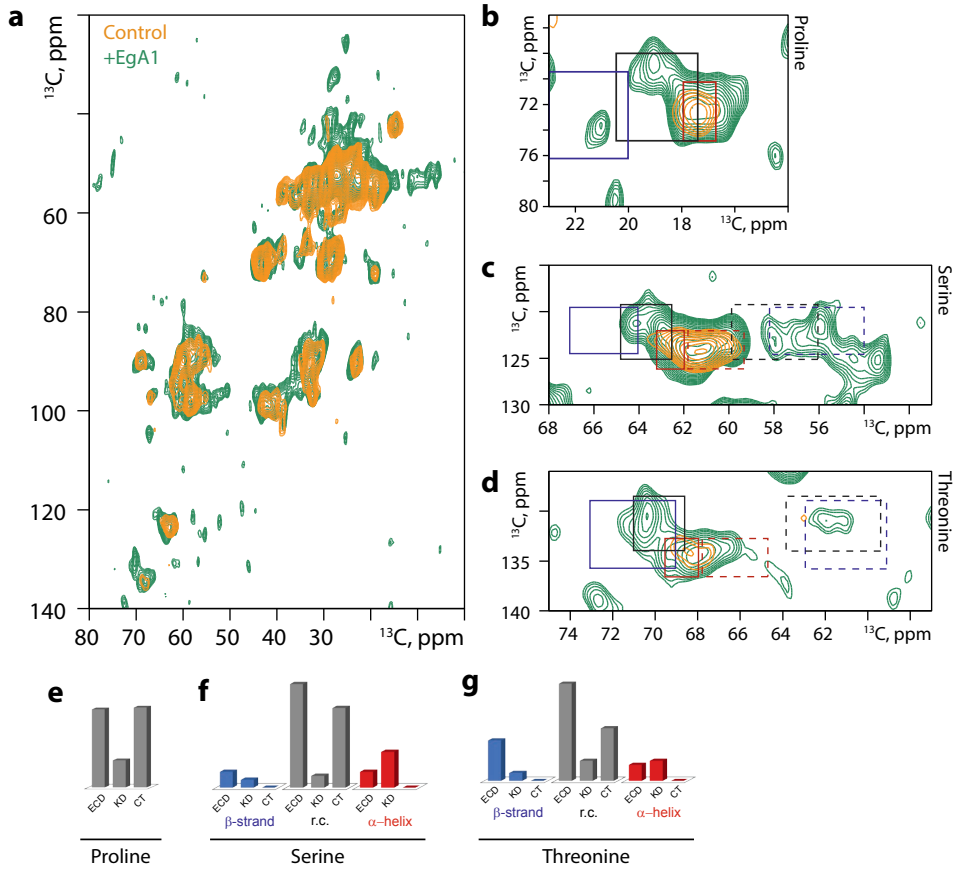


Figure 6 | EgA1-induced EGFR ECD rigidification drives HER2 ECD stabilization. a, two-dimensional DQSQ spectra recorded on fully labelled SKBR3 plasma membrane vesicles in the presence (green) of absence (orange) of EgA1. **b-d**, highlights of the spectrum presented in **a**, which contain secondary structure information for proline, serine and threonine respectively. **b,** Proline residues are enriched in the ECD and C-terminal tail (CT), suggesting that the addition of EgA1 leads to a rigidification of the ECD and/or CT. **c**, **d**, Serine and threonine residues also display information on the secondary structure elements as highlighted by the coloured boxes. Red boxes portray confidence intervals for α-helices (Cα solid box; Cβ dashed box), black boxes for random coil structures and the blue boxes for β-sheet resonances. Addition of EgA1 induces a general increase in random coil- and β-sheet confidence intervals which are secondary structure elements predominantly found in the HER2 ECD. **e-g**, Relative distribution of proline, serine and threonine residues in HER2 subdomains, in addition to the distribution of secondary structure information for serine and threonine residues. ssNMR spectra were recorded at 9 kHz MAS at ambient temperature.

HER2-overexpressing cell line SKBR3. HER2 was functional in cells and remained functional after the extraction of the plasma membrane. The expression of HER2 as compared to other ErbB family members was much higher, which enables a characterization of the dynamic profile of the full-length HER2 receptor under native conditions by ssNMR.

Our ssNMR data suggest that nonstimulated HER2 exhibits a dynamic behaviour where the ECD samples the conformational landscape, as previously observed for EGFR.

When the SKBR3 plasma membrane vesicles are not stimulated, the KD of HER2 is a rigid entity while the extracellular domain and C-terminal tail undergo significant motion leading to a reduction of NMR signal intensities After stimulation with EGF, we observe a stabilization of the Her2 ECD domain which would be compatible with a heterodimerzation of EGFR and HER2 across the entire surface of the plasma membrane, leading to an overall increase in phosphorylation levels. In this process, binding of EGF to DI and DIII of EGFR stabilizes the ECD which allow for the heterodimerization with HER2. In line with these data, a αEGFR nanobody EgA1 which has been shown to partially stabilize the EGFR ECD also induced HER2 phosphorylation. Moreover, addition of EgA1 to fully-labelled SKBR3 plasma membrane vesicles induced a full stabilization of HER2.

We propose a model where the induced rigidification after ligand binding is the driving force for HER2 activation. This would imply that heterodimerization of HER2 is a sequential process. First, unstimulated receptors are sampling different conformations. After binding of the ligand to EGFR, a the ECD is stabilized exposing rigid docking sites, which allow heterodimerization and HER2 phosphorylation. Such a EGFR:HER2 dimer appears not to be a stable complex as EGFR internalization occurs, while HER2 remains on the plasma membrane. The instability of dimer interfaces allows EGFR to be a seed for multiple HER2 complexes via kiss-and-run activation[40], thereby fully activating the much larger HER2 pool. This process has previously been shown to also been shown to involved in EGFR homodimerization and can quickly propagate dimerization across the plasma membrane[40] and might play a role in regulation of receptor functioning.

In contrast to previously thought, unstimulated HER2 exhibits a dynamic ECD which is stabilized after addition of EGF. Similar dynamics have been observed for the EGFR ECD and proven to be crucial for its functioning (see chapter 3). HER2 ECD motion might be required for HER2 to adapt to a range of docking partners, which is supported by the notion that Cetuximab, but not 7D12, is able to prevent EgA1-induced HER2 phosphorylation. Cetuximab and 7D12 have the same binding epitope, which suggests that Cetuximab prevents phosphorylation in a size-dependent manner rather than competing for a binding epitope on EGFR. Multiple ligands have been identified that bind to EGFR or other ErbB family members which stabilize a variety of structures that consequently also induce signalling pathways in unique manners. This would suggest a role for conformational sampling to allow for an ensemble of different structures to be stabilized upon ligand binding, which gives rise to a variety of responses. In support of this hypothesis is the conformational sampling which been observed for other types of receptors, where the stabilization of different structures affects receptor functioning, e.g. binding affinities. Conformational sampling might be a general feature of receptors which can be used by cells to regulate receptor functioning.

FUTURE PERSPECTIVES

Although our results provide valuable new insights in the molecular heterodimerization mechanism of HER2, there are still several open questions that remain to be answered. For example, our data shows that HER2 is more abundant on SKBR3 cells than EGFR and suggest that transient heterodimers are formed after EGF addition. However, it is still unknown how these receptors are distributed along the plasma membrane prior to- and after stimulation. For example, it would be possible that EGFR forms an oligomerization seed on one side of a clusters to which HER2 receptors bind, forming a "train" of receptors. Super-resolution microscopy techniques like gSTED could provide sufficient resolution to characterize the

heterodimerization process on a macroscopic level. Additionally to the heterodimerizaion process, it is also not clear what dimer interface forms the structural seed for the HER2 ECD to stabilize. Our preliminary data suggests that the front-to-front dimer interface is involved, however, additional control experiments should be included to confirm whether this is true and whether the back-to-back dimer interface is not involved.

MATERIALS AND METHODS

Cell culture. SKBR3- and A431 cells were cultured until 80-90% confluency in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and Penicillin/Streptomycin in an atmosphere containing 5% CO2 and 37°C.

Nanobody production. *E. coli* BL21C+ cells containing plasmid DNA for his-tagged EgA1 and 11A4 were grown overnight in LB as a preculture. On the morning of the production, overnight cultures were diluted 1/500 into auto-induction medium [1/50 of 5052 (0.5% glycerol, 0.05% glucose, 0.2% α-lactose), 1/20 NPS (100 mM NaHPO4, 25 mM $(NH_4)2SO_4$, 50 mM KH_2PO_4), 2mM $MgSO_4$, micronutrients and kanamycin in LB]. The next day, the bacteria were harvested and resuspended in PBS. After several freeze-thaw cycles, the periplasm fraction was collected. The nanobodies were purified using immobilized metal-affinity chromatography (IMAC) using the His-tag. A HisTrap HP Ni2+-Sepharose column was used to trap His-tag containing proteins. The nanobodies were reconsitutated at 1mg/mL concentration for storage at 4°C.

Mass spectrometry. Quantitative mass spectrometry analysis of protein levels was performed as described previously[18]. In short, SKBR3 plasma membrane vesicles and cells were resuspended in lysis buffer (50 mM ammonium bicarbonate, 1% SDC, 10 mM TCEP, 100mM TRIS, 40 mM chloroacetamide and complete protease inhibitor cocktail (Roche)) and boiled for 5 min. The peptides in the were digested overnight using LysC (1:75) and trypsin (1:50). The supernatant was desalted using C18 SepPak (Waters) cartridges, vacuum-dried and stored at -80°C for further analysis. Peptide mixtures were reconstituted in 10% formic acid and 1µg of protein digest of each sample was analyzed by nano-LC-MS/MS on an Orbitrap Q-Exactive plus (ThermoFisher Scientific, Bremen). The digest was trapped on an in-house made trap column (Reprosil pur C18, dr maisch, 100 µm x 2 cm, 3 µm) by loading with buffer A (A: 0.1% formic acid) and separated on an analytical column (Poroshell 120 EC C18, Agilent Technologies, 50 µm x 50 cm, 2.7 µm) using linear gradient from 13% to 40% buffer B (B: 0.1% formic acid, 80% ACN). During each scan cycle, the 10 most intense peptide precursors were selected for higher-energy collisional dissociation (HCD). Raw data files were processed with MaxQuant version 1.5.3.30. For relative quantification, intensities of all unique and razor peptides of a protein were summed up[41].

Phosphorylation Assay. SKBR3 cells were grown in a 6-well culture plate to 70% confluency before the medium was replaced to DMEM without FBS the day before the experiment. The cells were serum-starved a second time for 30 minutes prior to the experiment. Next, the cells were washed with binding buffer (DMEM –FBS buffered with 10 mM HEPES, pH 7.4) after which the cells were preincubated for 1 hour at 4°C with the respective nanobodies to allow maximal binding of the nanobodies. EGF (33 nM, R&D Systems) was added to the cells or plasma membrane vesicles to induce phosphorylation of HER2 for several time periods. The cells were washed twice with ice-cold phosphate-buffered saline (PBS) followed by cell lysis using a RIPA buffer (150 mM sodium chloride, 1.0% Triton X-100, 0.1% SDS, 50 mM Tris pH 8.0 and a protease inhibitor cocktail (Complete, Roche)). Proteins were separated by SDS-PAGE followed by western blotting onto PVDF-membrane. The membrane was incubated with Mouse-α-HER2 (Cell Signalling Technology, Danvers, Massachusetts; 44E7), Rabbit-α-phosphoHER2 Y1248 (Cell Signalling Technology, Danvers, Massachusetts; #2247) and Mouse-α-Actin overnight at 4°C. The antibody staining was visualized using Goat-α-Rabbit IRDye700 (Li-COR) and Goat-α-Mouse IRDye800 (Li-COR) and detected on an Odyssey imaging system (Li-COR).

Binding Assay. SKBR3- and A431 cells were grown in a 96-well Nunclon culture plate (ThermoFisher) to 80-90% confluency. The cells were washed with binding buffer (DMEM–FBS buffered with 10 mM HEPES, pH 7.4) after

which the cells were incubated for 2 hours at 4°C with a serial dilution of the respective nanobodies. The cells were washed twice with ice-cold PBS before they were fixed using a 4% paraformaldehyde solution (PFA) for 15 minutes at room temperature. For nanobody detection, the cells were incubated for 1 hour at room temperature with Rabbit-α-VHH (Creative BioLabs) followed by an incubation of 1 hour at room temperature of Goat-α-Rabbit IRDye800 (Li-COR). The 96-well plate was read using an Odessey imaging system (Li-COR) and wells were quantified using the Odessey software (Li-COR) and ImageJ.

Confocal Microscopy. A431 cells were seeded the day before the experiment at a concentration of 100.000 cells/well onto 12 mm coverslips in a 12-well culture plate. The cells were preincubated with 11A4 (100nM) for 30 minutes at 37°C followed by the addition of EGF-A488 (33 nM) for another 15 minutes at 37°C. The cells were washed twice with ice-cold PBS and fixed using a 4% PFA solution. 11A4 was indirectly stained by an overnight incubation at 4°C with Rabbit-α-VHH (Creative BioLabs) followed by a staining with Goat-α-RabbitA555. Confocal imaging was performed on a ZEISS LSM700 confocal microscope at 63x magnification.

ssNMR sample preparation. For a detailed description of this procedure see Kaplan et al[18]. In short, SKBR3 cells are grown to 80-90% on DMEM supplemented with an algal mixture of [^{13}C, ^{15}N] isotope-labelled amino acids (Cortecnet), 10% dialysed FBS, stable glutamine (2 mM) and Penicillin/Streptomycin in an atmosphere of 5% CO_2 and 37°C. The cells were resuspended in homogenizer buffer (10 mM Tris pH 7.4, 250 mM sucrose, 1 mM EDTA, Phosphatase inhibitors (100 μM sodium-orthovanadate) and protease inhibitors (Complete, Roche) before they were vesiculated by passing the cells through a syringe (21G, 1.5; 0.2x40 mm). Cellular debris was removed by centrifugation rounds at 1000x g until no visible pellet was found. The vesicles were collected by centrifuging at 150.000x g for 30 minutes at 4°C.

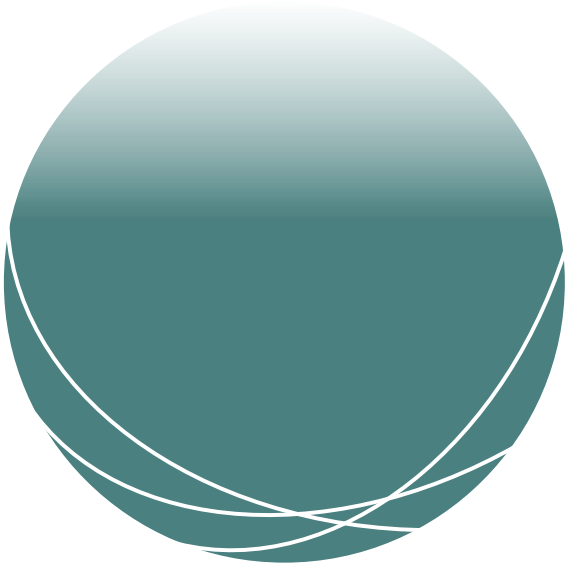
Solid-state NMR experiments. ssNMR experiments were performed using a standard-bore 700 MHz system (Bruker Biospin). Fully-labelled vesicle pellets were used to fill a 3.2 mm rotor. The samples were spun at 9 kHz MAS at a set temperature of 283K. 2D DQSQ spectra were recorded with a CP time of 500 μs and a SPC5 mixing time of 2.3 ms. In total, 110 t_1 points were recorded with a spectral window of 46656.176 Hz. The spectrum was processed using an EM function, with line broadening of 100 Hz for both t_1 and t_2 and 1k and 2k zero filling for t_1 and t_2 respectively. Differences in sample intensity due to variance in sample quantity were corrected for using 90° direct excitation of ^{13}C.

REFERENCES

1. Pinkas-Kramarski, R. et al. Diversification of Neu differentiation factor and epidermal growth factor signaling by combinatorial receptor interactions. EMBO J. (1996). doi:10.1002/j.1460-2075.1996.tb00603.x
2. Arteaga, C. L. & Engelman, J. A. ERBB receptors: From oncogene discovery to basic science to mechanism-based cancer therapeutics. Cancer Cell (2014). doi:10.1016/j.ccr.2014.02.025
3. Yarden, Y. The EGFR family and its ligands in human cancer. signalling mechanisms and therapeutic opportunities. Eur. J. Cancer (2001). doi:10.1016/S0959-8049(01)00230-1
4. Wahler, J. & Suh, N. Targeting HER2 Positive Breast Cancer with Chemopreventive Agents. Current Pharmacology Reports (2015). doi:10.1007/s40495-015-0040-z
5. Slamon, D. J. et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science (80-.). (1989). doi:10.1126/science.2470152
6. Citri, A., Skaria, K. B. & Yarden, Y. The deaf and the dumb: The biology of ErbB-2 and ErbB-3. in The EGF Receptor Family: Biologic Mechanisms and Role in Cancer (2003). doi:10.1016/B978-012160281-9/50005-0
7. Spears, M. et al. In situ detection of HER2:HER2 and HER2:HER3 protein-protein interactions demonstrates prognostic significance in early breast cancer. Breast Cancer Res. Treat. (2012). doi:10.1007/s10549-011-1606-z
8. Pernas, S. & Tolaney, S. M. HER2-positive breast cancer: new therapeutic frontiers and overcoming resistance. Therapeutic Advances in Medical Oncology (2019). doi:10.1177/1758835919833519
9. Ogiso, H. et al. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell (2002). doi:S0092867402009637 [pii]
10. Ferguson, K. M. et al. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol. Cell (2003). doi:10.1016/S1097-2765(03)00047-9
11. Cho, H. S. et al. Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature (2003). doi:10.1038/nature01392
12. Cho, H. S. & Leahy, D. J. Structure of the extracellular region of HER3 reveals an interdomain tether. Science (80-.). (2002). doi:10.1126/science.1074611
13. Bouyain, S., Longo, P. A., Li, S., Ferguson, K. M. & Leahy, D. J. The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. Proc. Natl. Acad. Sci. U. S. A. (2005). doi:10.1073/pnas.0507591102
14. Lemmon, M. A. Ligand-induced ErbB receptor dimerization. Experimental Cell Research (2009). doi:10.1016/j.yexcr.2008.10.024
15. Franco-Gonzalez, J. F., Ramos, J., Cruz, V. L. & Martinez-Salazar, J. Exploring the dynamics and interaction of a full ErbB2 receptor and Trastuzumab-Fab antibody in a lipid bilayer model using Martini coarse-grained force field. J. Comput. Aided. Mol. Des. (2014). doi:10.1007/s10822-014-9787-2
16. Alvarado, D., Klein, D. E. & Lemmon, M. A. Structural Basis for negative cooperativity in growth factor binding to an EGF receptor. Cell (2010). doi:10.1016/j.cell.2010.07.015
17. Arkhipov, A., Shan, Y., Kim, E. T., Dror, R. O. & Shaw, D. E. Her2 activation mechanism reflects evolutionary preservation of asymmetric ectodomain dimers in the human EGFR family. Elife (2013). doi:10.7554/eLife.00708
18. Kaplan, M. et al. EGFR Dynamics Change during Activation in Native Membranes as Revealed by NMR. Cell (2016). doi:10.1016/j.cell.2016.10.038
19. Freed, D. M. et al. EGFR Ligands Differentially Stabilize Receptor Dimers to Specify Signaling Kinetics. Cell (2017). doi:10.1016/j.cell.2017.09.017
20. Coskun, Ü., Grzybek, M., Drechsel, D. & Simons, K. Regulation of human EGF receptor by lipids. Proc. Natl. Acad. Sci. U. S. A. (2011). doi:10.1073/pnas.1105666108
21. Miljan, E. A. & Bremer, E. G. Regulation of growth factor receptors by gangliosides. Science's STKE : signal transduction knowledge environment (2002). doi:10.1126/stke.2002.160.re15
22. Liu, Y. C. et al. Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. Proc. Natl. Acad. Sci. U.

S. A. (2011). doi:10.1073/pnas.1107385108

23. Kaplan, M. et al. Probing a cell-embedded megadalton protein complex by DNP-supported solid-state NMR. Nat. Methods (2015). doi:10.1038/nmeth.3406
24. Narasimhan, S. et al. DNP-Supported Solid-State NMR Spectroscopy of Proteins Inside Mammalian Cells. Angew. Chemie Int. Ed. (2019). doi:10.1002/anie.201908849
25. Renault, M. et al. Cellular solid-state nuclear magnetic resonance spectroscopy. Proc. Natl. Acad. Sci. U. S. A. (2012). doi:10.1073/pnas.1116478109
26. Albert, B. J. et al. Dynamic Nuclear Polarization Nuclear Magnetic Resonance in Human Cells Using Fluorescent Polarizing Agents. Biochemistry (2018). doi:10.1021/acs.biochem.8b00257
27. Etzkorn, M. et al. Complex Formation and Light Activation in Membrane-Embedded Sensory Rhodopsin II as Seen by Solid-State NMR Spectroscopy. Structure (2010). doi:10.1016/j.str.2010.01.011
28. Schneider, R. et al. Probing molecular motion by double-quantum (13C,13C) solid-state NMR spectroscopy: Application to ubiquitin. J. Am. Chem. Soc. (2010). doi:10.1021/ja906283h
29. Hong, M., Zhang, Y. & Hu, F. Membrane Protein Structure and Dynamics from NMR Spectroscopy. Annu. Rev. Phys. Chem. (2012). doi:10.1146/annurev-physchem-032511-143731
30. Damman, R. et al. Atomic-level insight into mRNA processing bodies by combining solid and solution-state NMR spectroscopy. Nat. Commun. (2019). doi:10.1038/s41467-019-12402-3
31. Sarup, J. C. et al. Characterization of an anti-p185HER2 monoclonal antibody that stimulates receptor function and inhibits tumor cell growth. Growth Regul. (1991).
32. Bocharov, E. V. et al. Alternative packing of EGFR transmembrane domain suggests that protein-lipid interactions underlie signal conduction across membrane. Biochim. Biophys. Acta - Biomembr. (2016). doi:10.1016/j.bbamem.2016.02.023
33. Bocharov, E. V. et al. Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state. J. Biol. Chem. (2008). doi:10.1074/jbc.M709202200
34. Aertgeerts, K. et al. Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of HER2 protein. J. Biol. Chem. (2011). doi:10.1074/jbc.M110.206193
35. Alvarado, D., Klein, D. E. & Lemmon, M. A. ErbB2 resembles an autoinhibited invertebrate epidermal growth factor receptor. Nature (2009). doi:10.1038/nature08297
36. Schmitz, K. R., Bagchi, A., Roovers, R. C., Van Bergen En Henegouwen, P. M. P. & Ferguson, K. M. Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. Structure (2013). doi:10.1016/j.str.2013.05.008
37. Kijanka, M. et al. Rapid optical imaging of human breast tumour xenografts using anti-HER2 VHHs site-directly conjugated to IRDye 800CW for image-guided surgery. Eur. J. Nucl. Med. Mol. Imaging (2013). doi:10.1007/s00259-013-2471-2
38. Kijanka, M. et al. A novel immuno-gold labeling protocol for nanobody-based detection of HER2 in breast cancer cells using immuno-electron microscopy. J. Struct. Biol. (2017). doi:10.1016/j.jsb.2017.05.008
39. Needham, S. R. et al. EGFR oligomerization organizes kinase-active dimers into competent signalling platforms. Nat. Commun. (2016). doi:10.1038/ncomms13307
40. Verveer, P. J., Wouters, F. S., Reynolds, A. R. & Bastiaens, P. I. H. Quantitative imaging of lateral ErbB1 receptor signal propagation in the plasma membrane. Science (80-.). (2000). doi:10.1126/science.290.5496.1567
41. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. (2008). doi:10.1038/nbt.1511



Chapter 5

Atomic-level insight into mRNA processing bodies by combining solid and solution-state NMR spectroscopy

Reinier Damman[1], Stefan Schütz[2], Yanzhang Luo[1], Markus Weingarth[1], Remco Sprangers[2], and Marc Baldus[1]

[1] NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands.
[2] Department of Biophysics I, University of Regensburg, 93053, Regensburg, Germany.

Nat. Commun. ***10****, 4536 (2019); doi:10.1038/s41467-019-12402-3*

ABSTRACT

Liquid–liquid phase separation is increasingly recognized as a process involved in cellular organization. Thus far, a detailed structural characterization of this intrinsically heterogeneous process has been challenging. Here we combine solid- and solution-state NMR spectroscopy to obtain atomic-level insights into the assembly and maturation of cytoplasmic processing bodies that contain mRNA as well as enzymes involved in mRNA degradation. In detail, we have studied the enhancer of decapping 3 (Edc3) protein that is a central hub for processing body formation in yeast. Our results reveal that Edc3 domains exhibit diverse levels of structural organization and dynamics after liquid-liquid phase separation. In addition, we find that interactions between the different Edc3 domains and between Edc3 and RNA in solution are largely preserved in the condensed protein state, allowing processing bodies to rapidly form and dissociate upon small alterations in the cellular environment.

INTRODUCTION

Increasing evidence shows that cells develop intracellular protein-rich organelles that grow and fuse, allowing certain molecules to become locally enriched while excluding others[1]. Intracellular liquid–liquid phase separation (LLPS) processes seem to be a common route to achieving spatial organization of cellular components into dynamic, membrane-less compartments. While physical models have helped to understand the formation of membrane-less protein compartments[2], studying such networks at atomic resolution has been challenging due to the their intrinsic mobility and heterogeneous nature[3,4] and because genetic approaches in a cellular context are complicated by the highly redundant nature of the clustering process.

NMR spectroscopy has been shown to provide unique structural insights into heterogeneous and dynamical systems at atomic resolution[5,6,7,8]. Previously, NMR in pure solutions has been used to decipher molecular interactions among LLPS components under in vitro conditions using titration studies[9,10] and determined 3D structures of fast tumbling folded subdomains[11]. However, studying the structural organization of liquid-like droplets using solution-state NMR methods is complicated by the reduced diffusion[12] of the liquid-liquid phase separated molecules.

Hence, direct insights into the structural and dynamical organization of liquid-like droplets has been limited[12]. Instead, solid-state NMR (ssNMR) has been employed to study formation[13,14] and structural organization[13,15,16,17,18,19] of protein hydrogels and fibrillar assemblies that may be closely related to LLPS[3,20]. For example, we have previously used ssNMR to investigate the hydrogel state of the FG repeat domain of the nucleoporin Nsp1p, revealing that transient amyloid-like β-sheet interactions among NTQS-rich protein regions are responsible for gelation and network formation[13,16]. Such on-pathway amyloid interactions were also proposed for hydrogel-forming peptides that contain GSY-amino acid rich stretches and, in isolation, form crystalline needles reminiscent of amyloid fibrils21. Indeed, ssNMR revealed characteristic β -strand arrangements in fibrils formed by the low complexity domain of the FUS RNA-binding protein[18]. A relationship between LLPS and fibril-like protein states and their relevance for human disease has also been examined in stress granules composed of RNA-binding proteins (RBPs) and RNA[22] and, more recently, for huntingtin exon1 assemblies in mammalian cells and yeast[23]. While LLPS and fibrillization may represent two mechanistically distinct processes, models that connect these two processes have been developed[20,22,24.] For example, Molliex et al.[22] have depicted the relationship between phase separation, fibrillization, and pathological inclusions by three steps with an intermediate maturation regime where two phases coexist in the granule state, which ultimately connects LLPS to disease. However, the structural relationship between liquid-liquid phase separation and maturation has remained elusive.

In the following, we combined solid- and solution-state NMR experiments to directly study the LLPS and maturation of processing bodies (P-bodies) that are dynamic cytoplasmic ribonucleoprotein (RNP) granules containing proteins that are involved in translational repression and mRNA degradation[25,26,27,28]. Their main constituents are mRNA and the mRNA decay machinery, including the Dcp1:Dcp2 mRNA decapping complex, the RNA helicase Dhh1, the Pat–Lsm1–7 complex, the exonuclease Xrn1 and the Edc3[25,29]. Edc3 has emerged as a central hub for processing body formation in yeast. The 50 kDa protein (Fig. 1a) comprises an N-terminal LSm domain that directly interacts with helical leucine-rich motifs (HLMs) in the P-body proteins Dcp2[30] and Pdc1[9]. C-terminal to the LSm domain, Edc3 contains a 120 amino acid long intrinsically disordered region (IDR). This IDR

is an interaction hub for the RNA helicase Dhh1[31], RNA and the C-terminal Edc3 YjeF_N dimerization domain[25].

Previously, we showed that it is possible to reconstitute processing body-like assemblies *in vitro* using purified components[9,25]. These Edc3-containing *in vitro* foci undergo a maturation process, which over time results in the irreversible formation of an insoluble gel-like state (Figure 1b). Based on solution-state NMR and phase separation diagrams, we identified key interactions in this sequential two-step process and found that interactions between the Edc3 LSm domain and HLMs in Dcp2 and interactions between the Edc3 IDR and RNA are required for LLPS. Recently, we noticed that the isolated Edc3 protein is also able to undergo LLPS and we have shown that the IDR and the YjeF_N domain are key for this process. Moreover, the ability of Edc3 to mediate phase transitions is significantly enhanced by RNA which can bridge between Edc3 protomers through interactions with the Edc3 IDR[25]. Solution-state NMR spectroscopy is ideally suited to study proteins and protein:RNA complexes before they undergo LLPS (Fig. 1b; left panel), however, it is less optimal to study full-length processing body components after phase separation (Fig. 1b, middle panel), due to the limited diffusion within the high-density phase and due to the fact that these *in vitro* processing bodies mature into a more solid-like phase that does not spontaneously dissolve into a liquid phase upon dilution[25] (Fig. 1b, right panel). To study these states, we here turned to ssNMR.

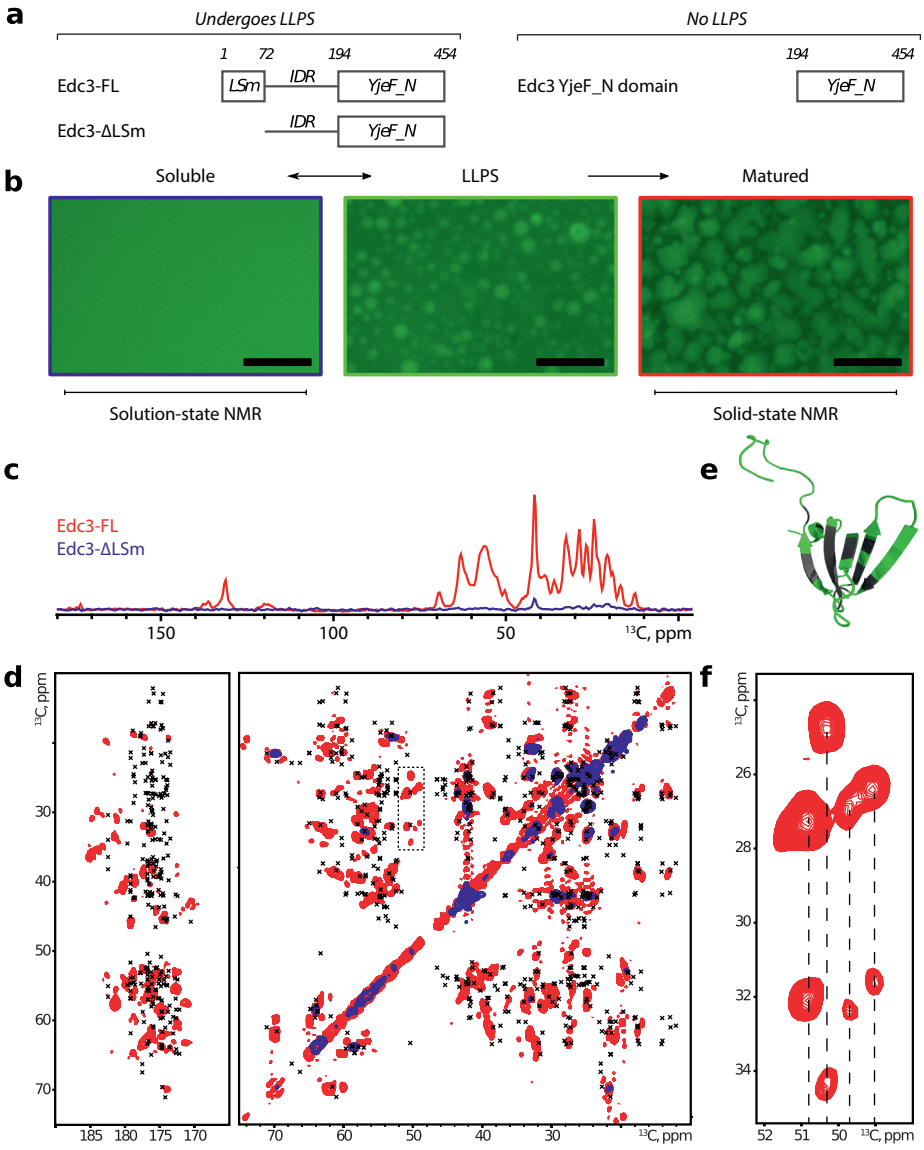
In the following, we conducted $^{13}C/^{15}N$- as well as ^{1}H-detected ssNMR experiments using a combination of scalar- and dipolar-based correlation methods that, as we have shown previously[13,16,32], provide a powerful tool to separate ssNMR protein signals of mobile and rigid protein regions. For ^{1}H-detected ssNMR experiments, we furthermore made use of proton dilution approaches[33,34] to increase spectral resolution under high-field/high-speed magic angle spinning (MAS) conditions. We applied our strategy to the matured state of full-length Edc3 protein as well as to the matured state of a Edc3 construct that lacks the N-terminal LSm domain (ΔLSm, Fig. 1a,b) representing the minimal Edc3 construct that can undergo LLPS[25]. In addition, we examined the C-terminal YjeF_N domain by both solid- and solution-state NMR and conducted ssNMR experiments on Edc3 preparations after the addition of RNA, which significantly enhances the LLPS process of Edc3.

Our studies reveal that the intra- and intermolecular contacts that are present between the Edc3 protein and RNA before LLPS are also existing in the matured granule state of *in vitro* processing bodies. These interactions are diverse and range from weak electrostatic to hydrophobic contacts and include folded protein domains as well as IDRs.

RESULTS

We initially examined the global dynamics of the matured state of full-length Edc3 (Edc3-FL; Fig. 1a,b) using scalar-based ssNMR experiments[13,16,32]. To this end, we recorded HC INEPT35 (Fig. 1c) and C-C TOBSY[36] (Fig. 1d) spectra that reveal regions of the protein that are highly mobile. These spectra indicate that at least one subdomain of the full-length Edc3 protein remains mobile, even in the matured (gel-like) state. As we previously showed that the Edc3

Figure 1 | The LSm domain of Edc3 remains mobile after LLPS. (**a**) Overview of the Edc3 constructs used in this study. Residue numbers are indicated on top. (**b**) Fluorescence microscopy images of the three states of Edc3 that we study here. The left panel (blue box) indicates the soluble form of the protein (or protein:RNA complex). This state is assessed using solution-state NMR spectroscopic methods. ➤



Under specific conditions (see below), the soluble Edc3 protein can undergo LLPS to form liquid-like foci (middle panel, green box). These foci mature over time to form a gel-like state that no longer dissolves upon dilution or by increasing the salt concentration (right panel, red box). The matured state of Edc3 is studied using solid-state NMR methods. The length of the scale bar in the images is 50 µm. (**c**) HC INEPT spectra highlighting dynamic regions in the full-length phase-separated Edc3 protein (Edc3-FL, red) and in the phase-separated Edc3 protein that lacks the N-terminal LSm domain (Edc3-ΔLSm, blue). (d) 2-dimensional C-C TOBSY spectra of the phase-separated Edc3-FL (red) and Edc3-ΔLSm (blue) proteins. The crosses indicate the solution chemical shifts of the LSm domain. (**e**) The NMR structure of the LSm domain color-coded for residues where the solution-state chemical shifts match with resonances in the C-C TOBSY spectrum (green) and for residues where the solution-state chemical shifts are not in a region where there are resonances observed in the C-C TOBSY spectrum (gray). (**f**) Zoom on the proline region highlighted by the black dashed box in panel c, indicating four unique proline spin systems.

LSm domain is not essential for the LLPS process of isolated Edc3[25] (*vide infra*), we speculated that this domain might remain flexible in the matured processing body state. To test this hypothesis, we repeated our scalar-based ssNMR experiments on matured processing bodies that were prepared from the Edc3 protein that lacks the LSm domain (Edc3-ΔLSm). In these spectra (Fig. 1c, d; blue) we observed a near complete reduction of the dynamics that were visible for the Edc3-FL protein. In particular, backbone correlations were strongly reduced in two-dimensional C-C TOBSY spectra of the matured states of the Edc3-ΔLSm proteins compared to Edc3-FL (Fig. 1d). Many of these disappearing correlations in Figure 1d matched with previous solution-state NMR assignments of the isolated LSm domain and the corresponding residues are indicated in Figure 1e in green. Deletion of the LSm domain removes two proline residues from the amino acid sequence. Yet, for Edc3-ΔLSm a total of four proline resonances disappeared (Fig. 1f), suggesting that the presence of the highly mobile LSm domain increases the flexibility of a part of the IDR. From these data, we conclude that the Edc3 LSm domain remains highly flexible in the matured state of the Edc3 protein. This notion also implies that the LSm domain is accessible for interactions with other processing body proteins and that it can thus engage in interactions with Dcp2, as we previously showed[30].

The C-terminal YjeF_N domain of Edc3 forms a strong dimer interface[37] that does not dissociate over time (Supplementary Figure 1) and that is essential for the LLPS behavior of Edc3[25]. Solution-state NMR ^{1}H-^{15}N TROSY spectra of the isolated and soluble YjeF_N domain display a well-folded protein (Fig. 2a, green). To assess if the fold of the YjeF_N domain changes upon maturation of the phase-separated Edc3 protein, we recorded ^{1}H-detected dipolar NH ssNMR spectra on the matured state of the Edc3-ΔLSm protein (Fig. 2a, gray). These spectra display Edc3 regions that are rigid on the time scale of our ssNMR experiment. As the Yjef_N domain comprised the majority of the Edc3-ΔLSm protein, we thus expected to observe mainly resonances from the dimerization domain. Indeed, we found a remarkable overall agreement between the NH correlation spectrum of the isolated Yjef_N domain in solution and the matured form of the Edc3-ΔLSm construct (Fig. 2a. Supplementary Figure 2, green v/s gray). The same agreement was observed for ^{1}H-^{13}C methyl TROSY[38] spectra recorded on the soluble YjeF_N domain as compared to a ^{1}H-detected dipolar CH ssNMR spectra recorded on the matured state of Edc3-ΔLSm (Fig. 2b, green v/s gray). Note that the solution-state NMR spectrum in Figure 2b only comprises signals of methyl groups from alanine-β, isoleucine-δ1, methionine-ε, valine-γ and leucine-δ positions, as the other carbon atoms and protons are NMR-inactive (see methods).

Even in the absence of residue-specific assignments, this strong resemblance of solution- and solid-state NMR data suggests that the fold of the YjeF_N domain is largely conserved in the matured state of the Edc3 protein without major structural rearrangements. This notion rules out a scenario where the Yjef_N domain forms β,-sheet-rich fibril-like structures after phase separation and maturation (Fig. 1b) and is further supported by the strong correlation between ssNMR spectra of the matured phase-separated Edc3-ΔLSm protein and the lyophilized Yjef_N domain (Supplementary Figure 3).

Despite the strong correlation between the ssNMR spectrum of the Edc3 protein in the phase-separated state and the solution-state NMR spectrum of the isolated Yjef_N domain, there are some notable differences. The major cause of this lies in the fact that the ssNMR spectrum was recorded on a protein construct that contains the IDR and the Yjef_N regions, whereas the solution-state NMR spectrum was obtained on a sample that

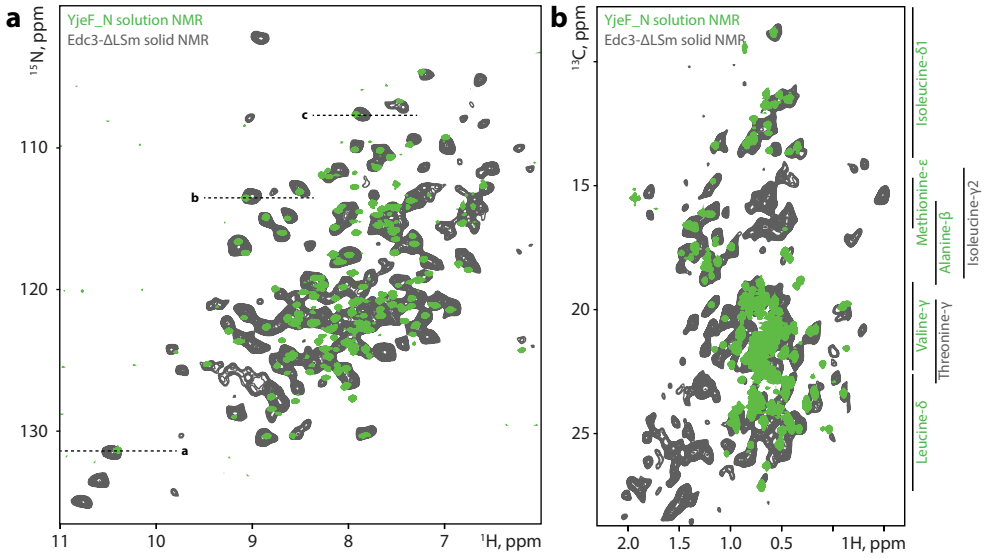


Figure 2 | The fold of the YjeF_N dimerization domain is conserved after LLPS. Comparison of dipolar NH ssNMR spectra recorded on Edc3-ΔLSm (gray) and of (1H-15N or methyl ^{1}H-^{13}C) TROSY-based spectra recorded with solution-state NMR on the YjeF_N domain (green). (**a**) ^{1}H-^{15}N region of the spectrum and (**b**) methyl region of the spectrum. The overall similarity between the spectra shows that the Edc3 Yjef_N domain is folded the same in solution and in the matured state. As a reference, average ^{13}C chemical shifts for methyl-containing residues are indicated next to ^{1}H-^{13}C spectrum (note that Thr-γ, Ile-δ2 methyl groups and CH_2 groups are not NMR-active in the solution-state samples). Dashed lines indicate the resonances used for 1D slices represented in Supplementary Figure 2.

only contains the Yjef_N domain. Note that it is not possible to record solution-state NMR spectra on a Edc3 construct that includes the IDR, as both Edc3-FL and Edc3-ΔLSm undergo LLPS and will transition into a matured state over time (see below), which is incompatible with solution-state NMR methods (Fig. 1b). Due to the presence of the IDR in the sample that was used to record the ssNMR spectra, additional NMR resonances are expected and observed (Fig 2a, gray v/s green spectrum). In addition, the IDR is known to interact with the Yjef_N domain in the phase-separated protein (Fig. 3a and see below). This will result in chemical shift perturbations of a set of the Yjef_N resonances and thus in differences between the solution and ssNMR spectra. In addition, as the YjeF_N domain for the solution-state NMR experiments was expressed in D_2O, not all backbone NH protons might have back-exchanged during the purification process, which could result in the absence of resonances in the solution-state ^{1}H-^{15}N NMR spectrum. Moreover, Ile-γ2 and Thr-γ methyl groups were not labeled for the solution- state NMR experiments and the corresponding resonances are thus not visible in the ^{1}H-^{13}C methyl TROSY spectrum. Finally, changes between the solution- and solid-state NMR spectra can result from differences in the measurement temperature (solid: 277 K, solution: 298K) and sample preparation. Based on our experience on globular proteins, we however judge that these temperature effects only cause minor changes in the NMR spectra. In summary, based on a comparison of Yjef_N spectra of the Edc3 protein in solution and in the matured state, we conclude that the structure of the Yjef_N domain is

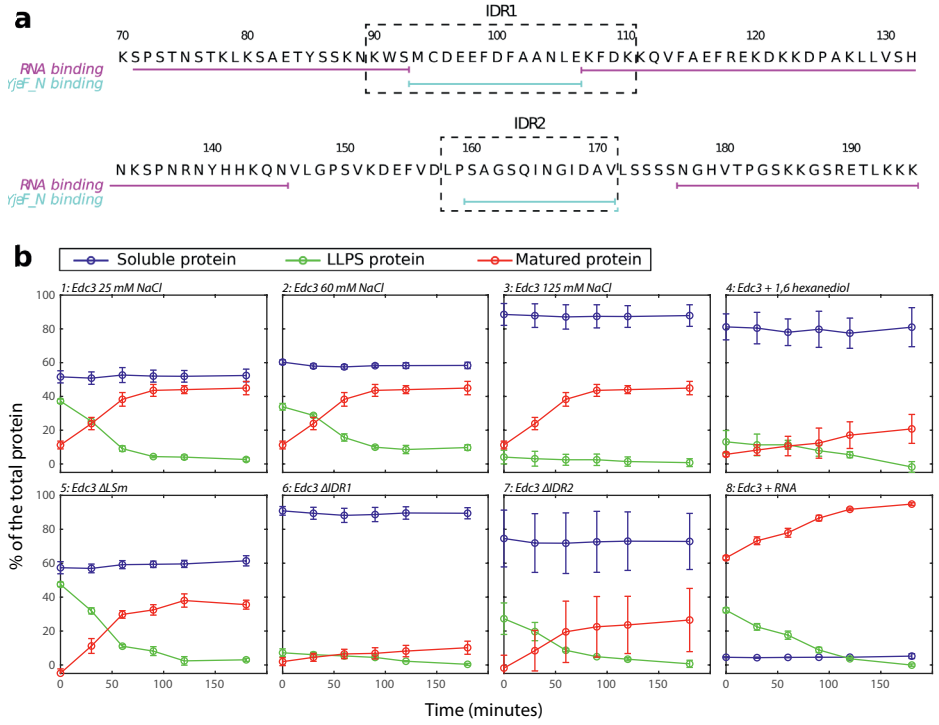


Figure 3 | Edc3 LLPS and maturation. (**a**) Edc3 IDR sequence with highlighted residues that interact with the YjeF_N domain (cyan lines) and RNA (magenta lines). (**b**) Under different conditions the fraction of total Edc3 protein that is in the soluble phase (blue lines; Fig. 1b, left panel), the LLPS phase (green lines; Fig 1b, middle panel) and in the matured phase (red lines; Fig 1b, right panel) were determined right after preparation of the solution and after 30, 60, 90, 120 and 180 minutes. Note that the decrease in the fraction of the protein that is in the LLPS phase correlates very well with the increase in the fraction of the protein that resides in the matured phase. This underscores that Edc3 maturation is a direct consequence of Edc3 LLPS (see Fig. 1b). Mean values and error bars (s.d.) are derived from three independent measurements.

largely maintained after LLPS.

Next, we assessed the importance of specific regions of Edc3 for LLPS and maturation. To that end, we performed in vitro experiments to determine to what degree various versions of Edc3 undergo LLPS and with what rate the condensed phase matures (Fig. 3b). In detail, we induced LLPS of Edc3-containing complexes and then measured over time the fraction of the total protein that is i) not phase-separated, ii) has undergone LLPS and iii) has matured into an insoluble gel (see methods). This data shows that the Edc3 protein at 25 mM NaCl (Fig. 3b, panel 1) fractionates such that ~50% of the total protein is soluble (blue line at timepoint 0), ~40% is in a condensed phase separation state (green line) and ~10% is in a matured state (red line). Over the course of an hour, the LLPS state matures and transforms into the insoluble matured state, such that ~50% of the protein is then in a soluble state (blue) and ~50% of the total protein is in the matured phase (red). Note that

the concentration of the soluble (blue) protein is constant at the critical concentration that is required for the formation of macroscopic phase separated droplets, whereas the protein that is in the LLPS phase (green) at time 0 is fully transformed into the matured gel-like state (red). At higher salt concentrations (Fig. 3b, panels 2 and 3), the critical concentration for LLPS increases and less protein participates into a LLPS state. As a result, the fraction of the protein that ends up in the matured state is lower at higher salt concentrations. From these data we conclude that electrostatic interactions play an important role in the LLPS process of the Edc3 proteins. Likewise, also upon addition of 5% 1,6-hexanediol, the phase separation propensity of Edc3 is reduced (panel 4), which indicates that hydrophobic interactions play a role in the Edc3 LLPS process. Interestingly, the addition of salt or 1,6-hexanediol not only increases the critical concentration for LLPS, it also appears to reduce the rate with which the LLPS protein matures. This indicates that the mode of the intermolecular interactions is similar in the LLPS and the matured state.

Based on the above setup, we addressed if the Edc3 LSm domain influences the LLPS and maturation process of the Edc3 protein. In agreement with the flexibility of the Edc3 LSm domain in the matured state of Edc3 (Fig. 1c-e), we observe that the removal of the Edc3 LSm domain has no significant effect of the LLPS extend and maturation rate (panel 5).

Previously, we have shown that the interaction between the Edc3 IDR and the Edc3 dimerization domain is important for the LLPS process of Edc3. In particular, we have identified two regions in the IDR that can interact with the Yjef_N domain[25]. These two regions comprise the amino acids K90-K111 (termed IDR1, Fig. 3a) and L159-L172 (termed IDR2, Fig. 3a), respectively. To assess the relative importance of these segments for the self-assembly process of Edc3, we performed LLPS and maturation assays with versions of the protein that lack either the first (IDR1, Fig. 3b, panel 6) or the second (IDR2, Fig 3b, panel 7) disordered segment that interacts with the Yjef_N domain (Fig. 3b). We found that removal of the first or second segment in the IDR strongly reduces the LLPS propensity of Edc3, where the first segment has a slightly stronger influence on LLPS than the second one. These results underscore the degeneracy in the interaction networks that are responsible for cellular phase transitions and identify IDR1 and IDR2 as important interaction sites for LLPS. To support this notion, we prepared [^{13}C, ^{14}N] and [^{12}C, ^{15}N] mixed-labeled Edc3-ΔLSm samples in the matured state and used these to probe intermolecular contacts between Edc3 dimers. Indeed, we observed intermolecular contacts between HN and HC atom pairs in NHHC experiments[39] verifying the presence of intermolecular contacts between two Edc3 dimers. These contacts most likely result from interactions of the IDR of one Edc3 dimer with the YjeF_N domain of another Edc3 dimer (Supplementary Figure 4).

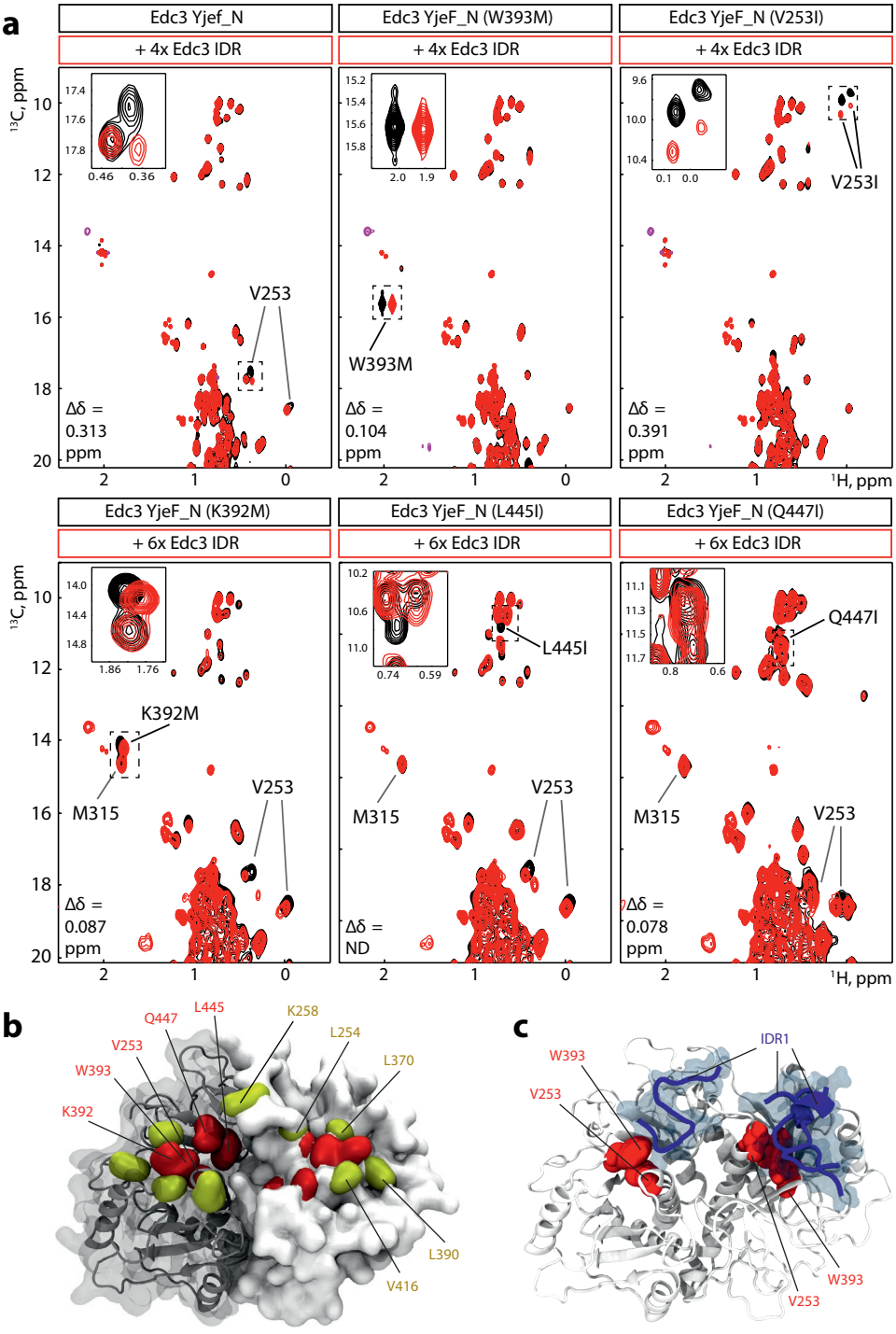
To address the interaction between the IDR and the Yjef_N domain with atomic detail, we turned to solution-state NMR techniques. In titration experiments we observed clear chemical shift perturbations in ^{1}H-^{15}N TROSY and ^{1}H-^{13}C methyl TROSY NMR spectra of the Yjef_N dimerization domain when the IDR is added in trans (Supplementary Figure 5a, b). To identify where the IDR interacts with the YjeF_N domain, we assigned a number of the resonances in the ^{1}H-^{15}N and ^{1}H-^{13}C spectra through a mutational approach (Fig. 4a). In particular, we were able to identify that W393 (based on ^{1}H-^{15}N data) and V253 (based on methyl TROSY data) in the dimerization domain are both part of the interface with the Edc3 IDR. As W393 and V253 are close in space on the surface of the Yjef_N domain they are part of the same binding pocket. To define additional residues that are part of this pocket we

took advantage of a slightly modified version of the methionine scanning approach[40,41]. In this method, methyl group-containing residues were introduced in the spatial vicinity to W393 and V253. These methyl groups were subsequently used as probes for the interaction of the Yjef_N domain with the IDR. Based on this strategy, we were able to identify that K392, L445 and Q447 are also part of the IDR binding pocket (Fig. 4a) as the introduced methyl groups experience chemical shift perturbations (CSPs) that are larger than 0.05 ppm. The introduced methyl groups did not alter the interaction mode between the Yjef_N domain and the IDR as the naturally occurring methyl groups of V253 underwent the same CSPs in the WT protein and the mutants. At the same time, we identified that K258, L254, L370, L390 and V416 are located outside the IDR binding interface (Supplementary Figure 5c), as these reporter methyl resonances are insensitive to the IDR, while the methyl groups of V253 remain effected by the IDR:Yjef_N interaction. In short, we find that the Edc3 IDR interacts with the dimerization domain at two specific pockets on the surface, one on each protomer (Fig 4b).

To shed further light on how the IDR interacts with the YjeF_N domain, we used atomistic molecular dynamics (MD) simulations (Fig. 4c). To this end, we used a homology model of the YjeF_N domain (see Materials & Methods) and performed six simulations of the complex formation between the YjeF_N domain in dimeric state and a short peptide that corresponds to IDR segment M93-E106 (IDR1). For each MD simulation, two 'IDR1 peptides' were randomly placed in the aqueous phase above the YjeF_Ndimer, shortly minimized and equilibrated with positions restraints, and the system was eventually freely evolved for 100 ns. Interestingly, in three of the six runs, we observed that peptides (in blue in Fig. 4c) bound the YjeF_N domain in direct proximity of V253 and W393 (in red), corroborating the solution-state NMR data. Peptides that bound in proximity of these residues also remained stably associated with the YjeF_N domain (Supplementary Figure 6).

Previously, we observed that RNA significantly enhances the LLPS tendency of the Edc3 protein as three distinct regions on the IDR can directly interact with RNA25 (Fig. 3a). Our data here shows that RNA significantly enhances the rate with which the phase-separated Edc3:RNA complex matures (Fig 3b; panel 8) as at the earliest timepoint, a large fraction of the protein already resides in the matured phase. Interestingly, for other proteins

Figure 4 | The IDR interacts directly with the YjeF_N domain. (**a**) The binding site of Edc3 IDR on the Edc3 Yjef_N domain was mapped using methyl TROSY NMR spectra. Specific residues on the Yjef_N surface were mutated into methyl group containing amino acids to add NMR-active probes. These mutant proteins were subsequently used in NMR titration experiments. Chemical shift perturbations of the introduced methyl groups ($\Delta\delta = \sqrt{(\Delta\delta_C^2 + \Delta\delta_H^2)}$) indicate if the corresponding residue is part of the Edc3 IDR binding site on the Edc3 Yjef_N domain and considered to be significant if $\Delta\delta > 0.05$ ppm (N.D.: due to spectral overlap it was not possible to determine the extent of the CSP). Spectra for methyl groups that we found to be outside the IDR binding site are shown in Supplementary Figure 4. Note that the naturally occurring methyl groups of Val 253 are shifted due to the W393M mutation and are thus not present at the WT resonance frequencies. In the V253I mutant protein, the methyl groups of V253 are not present and thus not detected in the spectra. In the W393M mutant, the V253 methyl groups are shifted to another spectra position and thus not readily identifiable. (**b**) The identified IDR binding site is highlighted onto the Yjef_N homology model. Residues of the YjeF_N domain that interact with the IDR (W393, V253, K392, L445 or Q447) are colored red while non-interacting residues (L370, V416, L254, K259 and L390) are colored in yellow. (**c**) MD simulations on the complex formation between the YjeF_N dimer (in white) and peptides corresponding to the first YjeF_N interacting residue stretch of the IDR M93-E106 (IDR 1). In the MD simulation, for three out of six replicas, the IDR peptides (in blue) directly interacted with residues V253 and W393 (in red) that were identified as YjeF_N-interacting hotspots in our solution-state NMR data. The figure shows the endpoints of the simulations after 100 ns. ►


a
Edc3 Yjef_N
+ 4x Edc3 IDR
Edc3 YjeF_N (W393M)
+ 4x Edc3 IDR
Edc3 YjeF_N (V253I)
+ 4x Edc3 IDR
^{13}C, ppm
V253
W393M
V253I
Δδ = 0.313 ppm
Δδ = 0.104 ppm
Δδ = 0.391 ppm
^{1}H, ppm
Edc3 YjeF_N (K392M)
+ 6x Edc3 IDR
Edc3 YjeF_N (L445I)
+ 6x Edc3 IDR
Edc3 YjeF_N (Q447I)
+ 6x Edc3 IDR
K392M
M315
V253
L445I
Q447I
Δδ = 0.087 ppm
Δδ = ND
Δδ = 0.078 ppm
b
Q447
L445
K258
V253
L254
W393
L370
K392
L390
V416
c
IDR1
W393
V253

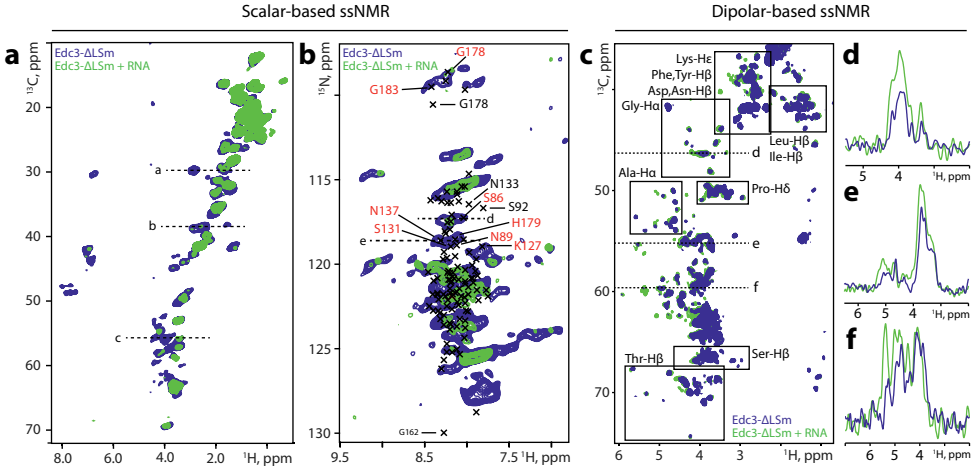


Figure 5 | RNA interacts with the IDR and results in reduced dynamics in the matured state. (**a,b**), Scalar-based CH and NH spectra that detect molecular elements with fast nano-second motions recorded on Edc3-ΔLSm in the absence (blue) and presence (green) of RNA, respectively. In panel **b**, the solution-state ^{1}H-^{15}N assignments of the IDR are indicated with black crosses. Putative interaction sites with RNA (red labels) or the YjeF_N domain (black labels) are indicated (See Fig. 3a). Dashed lines indicate 1D slices as represented in Supplementary Figure 7. (**c**) Dipolar CH spectra that detect rigid, immobile molecular entities were recorded on Edc3-ΔLSm in the absence (blue) and presence (green) of RNA. Several boxes indicate predicted chemical shifts for specific amino acids. (**d-f**), 1D slices of highlighted regions.

it has been shown that RNA prevents the formation of RNP granules[42], which highlights that small changes in the system can result in different LLPS and maturation behavior. To obtain insights into the molecular mechanism by which RNA is embedded into the matured state of phase-separated Edc3, we conducted ssNMR experiments on Edc3-ΔLSm protein in the presence of RNA. Interestingly, based on scalar-based CH and NH spectra we observed that the incorporation of RNA into the protein-rich Edc3 phase results in a significant loss in ssNMR signals, consistent with a reduction of the internal dynamics in the matured phase (Fig. 5a,b, Supplementary Figure 7; blue v/s green). In particular, we observe that many resonances that are typical for unstructured backbone- and side chains regions disappear upon interaction with RNA (Fig. 5b, green). This data thus directly reports on the interaction between RNA and the IDR which results in a loss of mobility of the IDR (Fig. 5b). Using assignments obtained from solution-state NMR25 (black crosses), the regions where signals disappear upon RNA incorporation (red labels) include S86, K127, S131, N137 and G183. This confirms our previous solution-state NMR analysis on the putative IDR:RNA interaction site (Fig. 3a), as these residues were identified to directly interact with RNA. Furthermore, many resonances of the IDR that are known to interact with the Yjef_N domain are not observed in these ssNMR spectra, as these amino acids are rigid both in the absence and in presence of the RNA.

To independently confirm the rigidification of the IDR upon interaction with RNA, we recorded dipolar CH experiments on the matured state of the Edc3-ΔLSm protein in the absence and presence of RNA. In those experiments, we observe that a number of peaks appears upon addition of RNA, confirming the notion that residues of the IDR are stabilized

by the RNA (Fig. 5c-f, green). The corresponding correlations (for example at 5.39 ppm ^{1}H and 59.9 ^{13}C ppm) are consistent with lysine side chain and serine/ threonine backbone ^{1}H-^{13}Cα peaks that are prominently present in the IDR segments that in solution interact with RNA (Supplementary Table 1). In conclusion, our solid-state NMR data suggest that RNA is recruited to the dynamic and unstructured IDR in Edc3, which results in decreased internal mobility.

DISCUSSION

LLPS is increasingly recognized as a cellular process that allows organizing cellular compartments across different spatial and temporal scales. A prominent example of such systems are RNP granules that contain proteins involved in translational repression and mRNA degradation[25].

Currently, the mechanisms by which LLPS modulates cellular function is not understood in sufficient detail[1,43]. One potential consequence of LLPS is that the catalytic activity of the embedded enzymes is effected, potentially through changes in the local concentrations or accessibility of substrates[25]. To obtain knowledge of these aspects, it is essential to gain atomic-level insights into the structural organization of LLPS proteins. Here, we demonstrated the power of combining solid- and solution-state NMR to study the process of phase separation events on the structural and dynamical level and at atomic scale.

Our results on the central processing body protein Edc3 reveal that the different domains exhibit different levels of structural organization and dynamics after liquid-liquid phase separation and a second transition into a matured phase (Fig 1b). Firstly, we find that the N-terminal LSm domain remains largely dynamic in the matured state, which leaves this domain accessible for contacts with additional processing body proteins including the Dcp1:Dcp2 mRNA decapping complex. Secondly, we find that the C-terminal YjeF_N domain largely retains its structure when the Edc3 protein undergoes a transition from the soluble to the matured phase separated state. This is in strong contrast to other proteins that have been observed to form insoluble amyloidic fibril structures after phase separation[18]. The minor structural changes in the fold of the Edc3 protein reflects the fact that proteins that reside in cellular processing bodies can rapidly dissociate from these foci.

Within the phase-separated state, the Edc3 Yjef_N domain interacts with the IDR of Edc3. This transient interaction is likely intermolecular as it increases the network of interactions leading to LLPS. Finally, we observe that the Edc3 IDR interacts with RNA in the matured state. These interactions involve distinct segments of the IDR and result in enhanced phase separation as the RNA extends and tightens the intermolecular interaction network.

Based on our data, we can construct a model for the formation of the matured state of the Edc3 protein (Fig. 6). In this model, the dimeric soluble Edc3 protein engages in a number of intermolecular interactions that involve RNA. These contacts result in the formation of an indefinite interaction network that drives the transition of the Edc3 protein from the dimeric soluble form into a dense LLPS protein state that then further matures into an insoluble state. Interestingly, our data clearly shows that the interactions between Edc3 and RNA that take place between the soluble components are, to a large degree, preserved in the matured phase separated state. The structural differences upon LLPS are thus small, which facilitates the reversibility of the process and necessitates only small cellular perturbations to induce phase separations (Fig. 6).

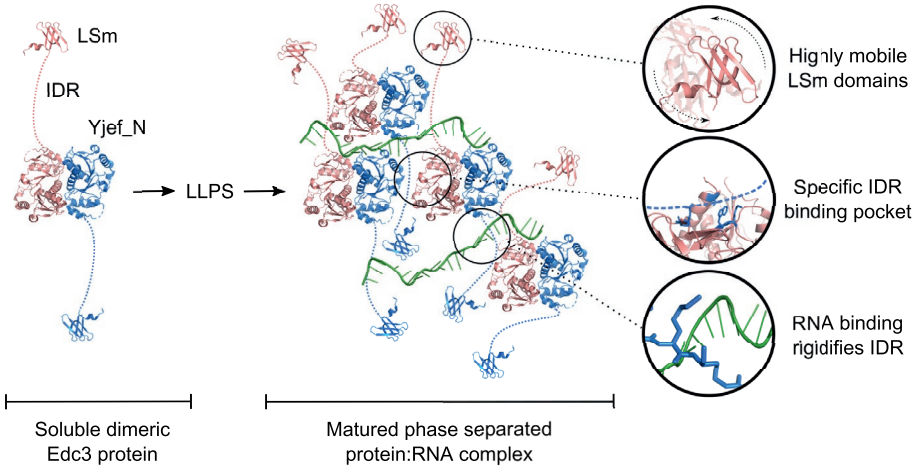


Figure 6 | Model of Edc3- and RNA-mediated processing body formation. The soluble dimeric Edc3 protein (blue and red) contains an N-terminal LSm domain that is linked through an IDR with the C-terminal dimeric Yjef_N domain (left). Intermolecular interactions between RNA (green), the IDR and the Yjef_N domains (solution-state NMR) result in the establishment of an infinite interaction network that leads to liquid-liquid phase separation of the RNP. The LLPS proteins mature over time to form a gel-like state (Fig. 1b). Within this matured phase of processing bodies, the LSm domains remain highly flexible (solid-state NMR) and can thus interact with other processing body proteins (top circle). Based on solution-state and solid-state NMR data, we have shown that the Edc3 YjeF_N dimerization domain remains dimeric and interacts specifically via a surface patch with an Edc3 IDR of another Edc3 dimer (middle circle). The RNA interacts with the Edc3 IDR, which results in a rigidification of the unstructured region (bottom circle).

Our work paves the way for future studies that address the exact structural changes that take place upon LLPS and those that result in the formation of the irreversible assemblies. Based on our data, these latter changes are small, but they will, nevertheless, have a significant effect on cellular homeostasis, as the maturation of cellular condensed protein phases needs to be prevented. We envision that our approach to study the structure and dynamics of LLPS complexes will be able to provide important insights in the current context and will also be applicable for other LLPS systems where irreversible assemblies are closely linked to fibril formation and disease.

ACKNOWLEDGEMENTS

This work was supported by the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007–2013), ERC Grant 616052 (RS). In addition, this study was funded by Netherlands Organization for Scientific Research (NWO) (grants 723.014.003 to M.W. and grants 700.26.121 and 700.10.443 to M.B.) and iNEXT (project number 653706), a Horizon 2020 program of the European Union. We thank David Meier and Olga Rudi for assisting in protein purification and Johanna Stöfl and Mira Schütz-Stoffregen for excellent technical assistance.

AUTHOR CONTRIBUTIONS

All authors designed the experiments; R.D,Y.L. and M.W. acquired solid state NMR data while S.S. acquired solution state NMR data. S.S. prepared samples; M.W. performed molecular dynamics simulations. All authors contributed to the analysis, interpretation of data and drafting of the manuscript.

MATERIALS & METHODS

Protein expression and purification. The genes for the full-length Edc3 protein, for Edc3 lacking the LSm domain (Edc3-ΔLSm) and for the Edc3 YjeF-N domain from *Schizosaccharomyces pombe* were cloned into modified pET vectors that carried an N-terminal TEV-cleavable His6- or MBP-tag (Supplementary Table 2). The gene for the IDR of Edc3 was cloned into a modified pET vector that carried an N-terminal TEV-cleavable His6-GST-tag. Point mutations and deletions were introduced into the genes using standard site-directed mutagenesis methods (Supplementary Table 3).

Escherichia coli BL21 (DE3) Codon Plus RIL cells (Stratagene) were transformed with the appropriate plasmid and grown at 37°C to an OD_{595} of 0.8 in LB. Overexpression was induced with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). After 18 hours at 20°C cells were pelleted by centrifugation and resuspended in buffer 1 (25 mM sodium phosphate, pH 7.4, 250 mM NaCl, 1 mM DTT) complemented with 10 mM imidazole, 5 mM $MgCl_2$, 0.1 % Triton X-100, lysozyme and 0.2 U/mL DNase I. Cells were lysed with ultrasound and insoluble cell debris were removed from the lysate by centrifugation. The supernatant was bound to equilibrated Ni-NTA resin, which was washed subsequently with buffer 1 complemented with 10 mM imidazole. Bound protein was eluted from the resin with buffer 1 complemented with 300 mM imidazole. Tobacco etch virus (TEV) protease was added to the elution fractions to cleave the purification tag from the target proteins.

The Edc3 IDR was dialyzed overnight at 4 °C into buffer 2 (25 mM HEPES, pH 8.0, 50 mM NaCl, 1 mM DTT) supplemented with 0.5 mM EDTA. The dialysate was applied to a cation exchange column (HiTrap 5mL SP FF, GE Healthcare) equilibrated with buffer 2. To separate the protein of interest from the cleaved GST-tag, a gradient from 0 to 35 % buffer 3 (25 mM HEPES, pH 8.0, 2 M NaCl, 1 mM DTT) was applied over a volume of 100 mL. Fractions containing the target protein were pooled and the buffer was exchanged to buffer 4 (25 mM HEPES, pH 7.3, 125 mM NaCl, 1 mM DTT) using centrifugal filters.

MBP-tagged Edc3 protein was purified from the cleared cell lysate (buffer 2) by anion exchange chromatography (HiTrap 5mL Q FF, GE Healthcare) using a gradient from 0 to 40 % buffer 3 over a volume of 100 mL. Fractions containing the target protein were pooled and the buffer was exchanged to buffer 4 by dialysis for subsequent size exclusion chromatography.

Proteins carrying only an N-terminal His6-tag were directly dialyzed into buffer 4 supplemented with 0.5 mM EDTA. Purification to homogeneity was achieved by size exclusion chromatography on HiLoad 16/600 Superdex 75 (Edc3 IDR, Edc3 YjeF_N) or Superdex 200 (Edc3-FL and Edc3-ΔLSm) columns (GE Healthcare) in buffer 4.

NMR-active proteins were obtained by overexpression in (50-100%) D_2O-based M9 minimal medium supplemented with 0.5 g/L $^{15}NH_4Cl$ as the sole nitrogen source and either 2 g/L $^{1}H^{13}C$-glucose for solid-state NMR samples or 4 g/L $^{1}H^{12}C$-glucose for solution-state NMR samples. For labeling of the Ile-δ1, Met-ε, Val-γ1/γ2 and Leu-δ1/δ2 methyl groups within an U-[$^{2}H^{12}C$]-background, the medium was supplemented with α-ketobutyrate (3-$^{2}H_2$-4-$^{13}CH_3$, 60 mg/L), methionine (methyl-$^{13}CH_3$; 100 mg/L), α-ketoisovalerate (3-^{2}H-3-(methyl-$^{13}CH_3$)-4-$^{13}CH_3$; 60 mg/L) one hour prior to induction. For simultaneous labeling of Ala-β methyl groups, 2-^{2}H-3-^{13}C-L-alanine was added to the medium 20 minutes prior to induction.

RNA in vitro transcription and purification. RNA in vitro transcription and purification was essentially carried out as described before[25]. In vitro transcription of a 30mer RNA was accomplished using in-house purified T7 polymerase[44] and two DNA primers as the DNA template (Supplementary Table 4). Transcribed RNA was purified under denaturing conditions using anion exchange chromatography on a DNAPac PA100 column (22 x

250 mm, Dionex) using a gradient from buffer 5 (20 mM Tris, pH 8,5 M urea) to buffer 6 (as buffer 5, supplemented with 2 M NaCl). The purified RNA was precipitated from the pooled fractions using 0.7 columes of isopropanol, followed by desalting on a PD10 column (GE Healthcare) and SpeedVac concentration. The obtained pure dry RNA product was resuspended at concentrations required for subsequent experiments. The quality of the RNA was assessed using urea-polyacrylamide gel electrophoresis (PAGE) in 1× TBE (89 mM Tris, pH 8.0, 89 mM boric acid, 2 mM EDTA). RNA was visualized by methylene blue staining.

Liquid-liquid phase separation assays. Different versions of the Edc3 protein (FL, ΔLSm, ΔIDR1 or ΔIDR2) in the presence or absence of salt, 5% 1,6-hexanediol or RNA were prepared at a final concentration of 150 μM by dilution from a non-phase-separated stock solution that was at a high salt concentration. Subsequently, two 8 μL aliquots of the sample were taken. To aliquot one 2 μL 2.5 M NaCl was added to dissolve the LLPS protein completely. This aliquot was cleared from any insoluble (matured) protein by centrifugation, after which the protein concentration of the supernatant was determined. This reading provides a measure for the protein that is not matured. The second aliquot was cleared from LLPS proteins and from matured proteins by centrifugation, after which the concentration of the Edc3 protein was determined. This reading provides a measure for the protein that is not phase separated or matured. The amount of the protein that resides in the matured phase was calculated based on the total protein concentration and the amount of the protein in the soluble + LLPS phases. These measurements were performed right after the preparation of the sample (0 minutes) and after 30, 60, 90, 120 and 180 minutes incubation at room temperature. For the Edc3:RNA mixture, the protein concentrations were determined via the Oregon green label that was attached to 20% of the total protein. All conditions contained 25 mM HEPES, pH 7.3 and 25 mM NaCl, unless indicated otherwise. The reported values are the average of three independent LLPS/maturations experiments (n=3 for all timepoints and for all conditions). The indicated error is the standard deviation of the independent measurements.

Solution-State NMR experiments. All solution-state NMR samples were in buffer 4 and contained 5% D_2O. NMR spectra were recorded at 298 K on Bruker AVIII-500 and AVIII-800 spectrometers equipped with room temperature and cryo probe-heads, respectively. NMR titration experiments were carried out with 0.05-0.1 mM ^{15}N- and ILVM methyl-labeled protein (Edc3 YjeF) and a 4- to 6-fold excess of unlabeled protein (Edc3 IDR). NMR spectra were processed using the NMRPipe/NMRDraw software suite[45]. Figures displaying NMR spectra and protein structures were produced using NMRview (onemoonscientific.com) and Pymol (pymol.org), respectively. Solution-state NMR parameters are given in an overview in Supplementary Table 5.

Solid-State NMR experiments. Phase-separated Edc3 samples for solid-state NMR measurements were obtained by lowering the salt concentration of concentrated protein from 125 mM to 25 mM by dilution with buffer 5 (25 mM HEPES, pH 7.3, 1 mM DTT). For samples to contain RNA, the protein was supplemented with a 2.5-fold molar excess of 30mer RNA and buffer 5. Protein and protein:RNA samples were incubated for 30 minutes at room temperature. Droplets were pelleted by centrifugation at 15.000 g for 15 minutes. The supernatant was discarded, and the gel was transferred to the solid-state NMR rotors.

Solid-state NMR experiments were conducted using 3.2 mm and 1.3 mm triple-resonance (^{1}H, ^{13}C, and ^{15}N) magic-angle-spinning (MAS) probe heads in static magnetic fields of 9.4, 16.4, 18.8 and 22.2 T, corresponding to proton resonance frequencies of 400, 700, 800 and 950 MHz, respectively. Scalar-based correlation experiments utilized HC INEPT[35] and 6 ms mixing C-C TOBSY[36] transfer steps. Samples were chilled to approximately 278 K sample temperature and spun at a MAS rate of 9 kHz. 1H-detected experiments were performed at a MAS rate of 60 kHz and a sample temperature of approximately 310 K. Dipolar-based sequences were used with cross-polarization (CP) steps with an amplitude-ramp of 80-100% on ^{1}H and 15 kHz PISSARRO[46] decoupling during detection periods. For J-based experiments, decoupled HSQC[47] sequences were used with a 4 and 3.35 ms INEPT transfer time for NH and CH transfers, respectively. NHHC experiments[39] were performed on [^{12}C,^{15}N]/[^{13}C,^{14}N]-mixed labeled samples in the matured state. Dipolar-based CP steps were used for magnetization transfer from ^{13}C to ^{1}H and from ^{1}H to ^{15}N. Proton-mixing times were from 0, 1, and 2 ms. Solid-state NMR parameters are given in an overview in

Supplementary Table 6.

Molecular Dynamics Simulations. MD simulations to probe the interaction between the YjeF_N domain and the Edc3 IDR were run with the g54a7 force field[48] and the GROMACS simulation package version 4.6.3.49. A 3D model of the dimeric YjeF_N domain was obtained using homology modelling[50] based on the structure of the human protein (PDB:3D3K [10.2210/pdb3D3K/pdb])[37]. A linear 14-meric peptide corresponding to the IDR stretch M93-E106 was built in Pymol (https://pymol.org/, version 1.3). For the starting state, peptides were placed within approximately 1 nm distance from the dimeric YjeF_N domain in a cubic aqueous box. The system was first energy minimized, afterwards equilibrated in an NVT ensemble for 100 ps with position restraints, then further equilibrated in a NPT ensemble for 100 ps with position restraints, and eventually freely evolved for 100 ns. In total, six replicates of 100 ns duration each with randomized initial peptide positions were run.

Contacts between IDR segment M93-E106 and the YjeF_N domain were quantified with the g-mindist tool of the GROMACS simulation package using a contact cut-off of 8 Å. Contacts were counted over the last 50 ns of the trajectory. Since the YjeF_N domain is symmetric, the contacts for both monomers were added together. The final analysis is the summation of the intermolecular IDR - YjeF_N domain contacts from three replicas.

REFERENCES

1. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. Nature Reviews Molecular Cell Biology 18, 285 (2017).
2. Brangwynne CP, Tompa P, Pappu RV. Polymer physics of intracellular phase transitions. Nature Physics 11, 899-904 (2015).
3. Wu H, Fuxreiter M. The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. Cell 165, 1055-1066 (2016).
4. Mitrea DM, et al. Methods for Physical Characterization of Phase-Separated Bodies and Membrane-less Organelles. Journal of Molecular Biology 430, 4773-4805 (2018).
5. Baldwin AJ, Kay LE. NMR spectroscopy brings invisible protein states into focus. Nature Chemical Biology 5, 808-814 (2009).
6. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. Nature Reviews Molecular Cell Biology 16, 18-29 (2015).
7. Hu KN, Tycko R. What can solid state NMR contribute to our understanding of protein folding? Biophysical Chemistry 151, 10-21 (2010).
8. Renault M, Cukkemane A, Baldus M. Solid-State NMR Spectroscopy on Complex Biomolecules. Angewandte Chemie International Edition 49, 8346-8357 (2010).
9. Fromm SA, Kamenz J, Noldeke ER, Neu A, Zocher G, Sprangers R. In Vitro Reconstitution of a Cellular Phase-Transition Process that Involves the mRNA Decapping Machinery. Angewandte Chemie-International Edition 53, 7354-7359 (2014).
10. Burke KA, Janke AM, Rhine CL, Fawzi NL. Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. Molecular Cell 60, 231-241 (2015).
11. Kuhn J, et al. The adaptor protein CIN85 assembles intracellular signaling clusters for B cell activation. Science Signaling 9, (2016).
12. Brady JP, et al. Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. Proceedings of the National Academy of Sciences of the United States of America 114, E8194-E8203 (2017).
13. Ader C, Frey S, Maas W, Schmidt HB, Gorlich D, Baldus M. Amyloid-like interactions within nucleoporin FG hydrogels. Proceedings of the National Academy of Sciences of the United States of America 107, 6281-6285 (2010).
14. Ackermann BE, Debelouchina GT. Heterochromatin Protein HP1 Gelation Dynamics Revealed by Solid-State NMR Spectroscopy. Angewandte Chemie-International Edition 58, 6300-6305 (2019).
15. Kennedy SB, deAzevedo ER, Petka WA, Russell TP, Tirrell DA, Hong M. Dynamic structure of a protein hydrogel: A solid-state NMR study. Macromolecules 34, 8675-8685 (2001).
16. Labokha AA, et al. Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. Embo Journal 32, 204-218 (2013).
17. Nagy-Smith K, Moore E, Schneider J, Tycko R. Molecular structure of monomorphic peptide fibrils within a kinetically trapped hydrogel network. Proceedings of the National Academy of Sciences of the United States of America 112, 9816-9821 (2015).
18. Murray DT, et al. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. Cell 171, 615-+ (2017).
19. Ward ME, Daniels MA, van Kappel EC, Maurice MM, Baldus M. Investigations of dynamic amyloid-Dlike structures of the Wnt signalling pathway by solid-state NMR. Chemical Communications 54, 3959-3962 (2018).
20. Xiang SH, et al. The LC Domain of hnRNPA2 Adopts Similar Conformations in Hydrogel Polymers, Liquid-like Droplets, and Nuclei. Cell 163, 829-839 (2015).
21. Hughes MP, et al. Atomic structures of low-complexity protein segments reveal kinked beta sheets that assemble networks. Science 359, 698-701 (2018).
22. Molliex A, et al. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. Cell 163, 123-133 (2015).
23. Peskett TR, Rau F, O'Driscoll J, Patani R, Lowe AR, Saibil HR. A Liquid to Solid Phase Transition Underlying Pathological Huntingtin Exon1

Aggregation. Molecular Cell 70, 588-+ (2018).
24. Lin Y, Protter DSW, Rosen MK, Parker R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Molecular Cell 60, 208-219 (2015).
25. Schutz S, Noldeke ER, Sprangers R. A synergistic network of interactions promotes the formation of in vitro processing bodies and protects mRNA against decapping. Nucleic Acids Research 45, 6911-6922 (2017).
26. Sheth U, Parker R. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science 300, 805-808 (2003).
27. van Dijk E, Cougot N, Meyer S, Babajko S, Wahle E, Seraphin B. Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. Embo Journal 21, 6915-6924 (2002).
28. Luo Y, Na ZK, Slavoff SA. P-Bodies: Composition, Properties, and Functions. Biochemistry 57, 2424-2431 (2018).
29. Decker CJ, Teixeira D, Parker R. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. Journal of Cell Biology 179, 437-449 (2007).
30. Fromm SA, et al. The structural basis of Edc3- and Scd6-mediated activation of the Dcp1:Dcp2 mRNA decapping complex. Embo Journal 31, 279-290 (2012).
31. Sharif H, Ozgur S, Sharma K, Basquin C, Urlaub H, Conti E. Structural analysis of the yeast Dhh1-Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions. Nucleic Acids Research 41, 8377-8390 (2013).
32. Andronesi OC, Becker S, Seidel K, Heise H, Young HS, Baldus M. Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy. Journal of the American Chemical Society 127, 12965-12974 (2005).
33. Mance D, et al. An Efficient Labelling Approach to Harness Backbone and Side-Chain Protons in H-1-Detected Solid-State NMR Spectroscopy. Angewandte Chemie-International Edition 54, 15799-15803 (2015).
34. Medeiros-Silva J, et al. H-1-Detected Solid-State NMR Studies of Water-Inaccessible Proteins InVitro and InSitu. Angewandte Chemie-International Edition 55, 13606-13610 (2016).
35. Morris GA, Freeman R. ENHANCEMENT OF NUCLEAR MAGNETIC-RESONANCE SIGNALS BY POLARIZATION TRANSFER. Journal of the American Chemical Society 101, 760-762 (1979).
36. Baldus M, Meier BH. Total correlation spectroscopy in the solid state. The use of scalar couplings to determine the through-bond connectivity. Journal of Magnetic Resonance Series A 121, 65-69 (1996).
37. Ling SHM, Decker CJ, Walsh MA, She MP, Parker R, Song HW. Crystal structure of human Edc3 and its functional implications. Molecular and Cellular Biology 28, 5965-5976 (2008).
38. Tugarinov V, Hwang PM, Ollerenshaw JE, Kay LE. Cross-correlated relaxation enhanced H-1-C-13 NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. Journal of the American Chemical Society 125, 10420-10428 (2003).
39. Lange A, Luca S, Baldus M. Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. Journal of the American Chemical Society 124, 9704-9705 (2002).
40. Stoffregen MC, Schwer MM, Renschler FA, Wiesner S. Methionine Scanning as an NMR Tool for Detecting and Analyzing Biomolecular Interaction Surfaces. Structure 20, 573-581 (2012).
41. Cvetkovic MA, Wurm JP, Audin MJ, Schutz S, Sprangers R. The Rrp4-exosome complex recruits and channels substrate RNA by a unique mechanism. Nature Chemical Biology 13, 522-+ (2017).
42. Maharana S, et al. RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science 360, 918-921 (2018).
43. Shin Y, Brangwynne CP. Liquid phase condensation in cell physiology and disease. Science 357, eaaf4382 (2017).
44. Fuchs AL, Neu A, Sprangers R. A general method for rapid and cost-efficient large-scale production of 5 ' capped RNA. Rna 22, 1454-1466 (2016).
45. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPIPE - A MULTIDIMENSIONAL SPECTRAL PROCESSING SYSTEM BASED ON UNIX PIPES. Journal of Biomolecular Nmr 6, 277-293 (1995).
46. Weingarth M, Bodenhausen G, Tekely P. Low-power

decoupling at high spinning frequencies in high static fields. Journal of Magnetic Resonance 199, 238-241 (2009).

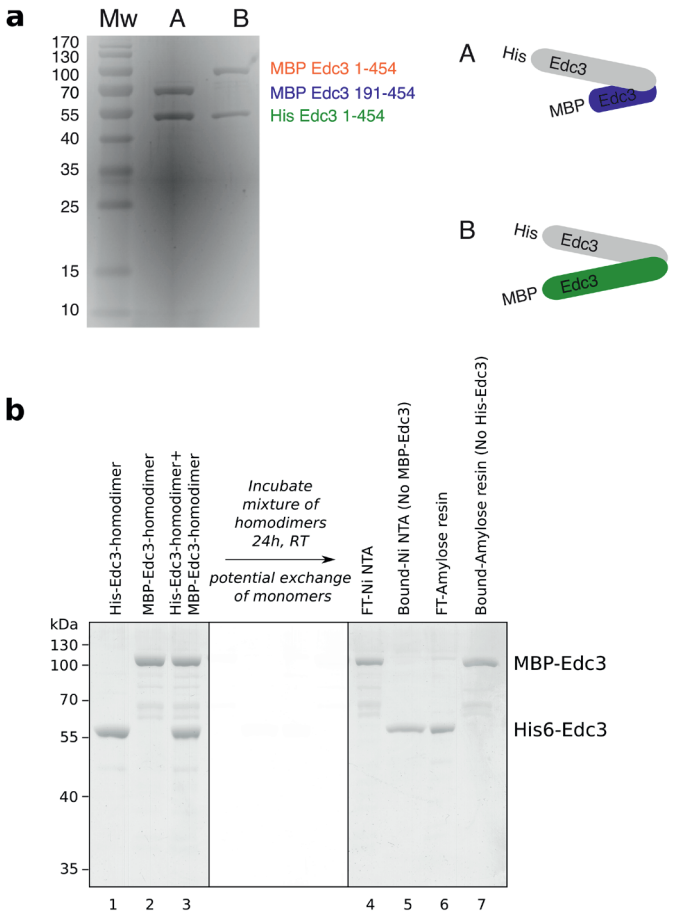
47. Bodenhausen G, Ruben DJ. NATURAL ABUNDANCE N-15 NMR BY ENHANCED HETERONUCLEAR SPECTROSCOPY. Chemical Physics Letters 69, 185-189 (1980).

48. Schmid N, et al. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. European Biophysics Journal with Biophysics Letters 40, 843-856 (2011).

49. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. Journal of Chemical Theory and Computation 4, 435-447 (2008).

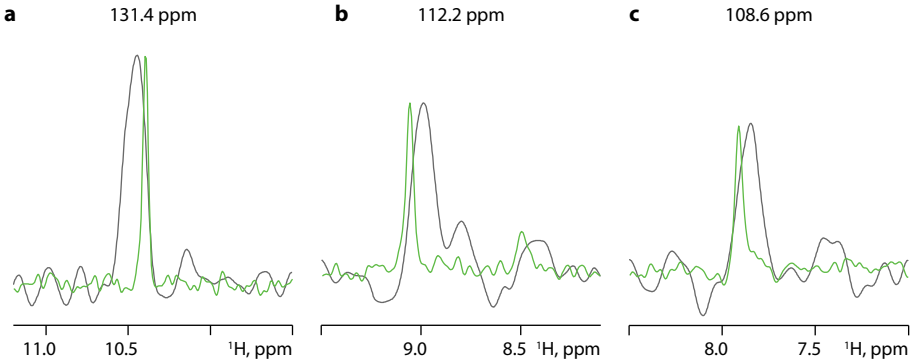
50. Sali A, Potterton L, Yuan F, Vanvlijmen H, Karplus M. EVALUATION OF COMPARATIVE PROTEIN MODELING BY MODELER. Proteins-Structure Function and Genetics 23, 318-326 (1995).

SUPPLEMENTARY INFORMATION

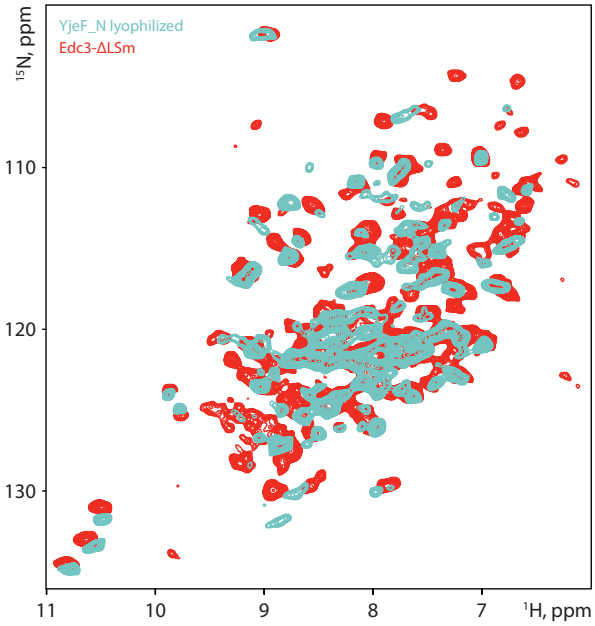


Supplementary Figure 1 | Edc3 forms a stable dimer in solution. (**a**) His-tagged full-length Edc3 was co-expressed in *E. coli* with either MBP-tagged Edc3 that contained Edc3 the Yjef_N domain (A) or with full length MBP-tagged full length Edc3 (B). The expressed proteins were purified in two steps. In step one all His-tagged Edc3 complexes (dimers) were selected for using Ni-NTA chromatography. In step two, all MBP-tagged Edc3 complexes (dimers) were selected for using amylose affinity chromatography. The Edc3 complexes that were purified this way thus contain at least one His-tag and at least on MBP-tag. The purified complexes were analyzed by SDS-PAGE and show a 1:1 ratio of a His-tagged and an MBP-tagged Edc3 monomer. This clearly shows that the Edc3 protein forms a stable dimer in solution through the Yjef_N domain. Mw: molecular weight marker in kDa. (**b**) To assess the long-term stability of the Edc3 dimers, His-Edc3 homodimers (lane 1) and MBP-Edc3 homodimers (lane 2) were separately expressed and purified. These dimers were subsequently mixed in a 1:1 ratio (lane 3). After 24 hours at room temperature, one half of the mixed dimers was applied to a Ni-NTA column. The FT of this column contains all Edc3 dimers that do not contain a His-Edc3 monomer (lane 4). The elution of this Ni-NTA column contains all Edc3 dimers that contain at least one His-Edc3 monomer (lane 5). These latter dimers include MBP-Edc3 homodimers as well as (in case exchange of dimers took place) MBP-Edc3:His-Edc3 heterodimers. These mixed dimers are

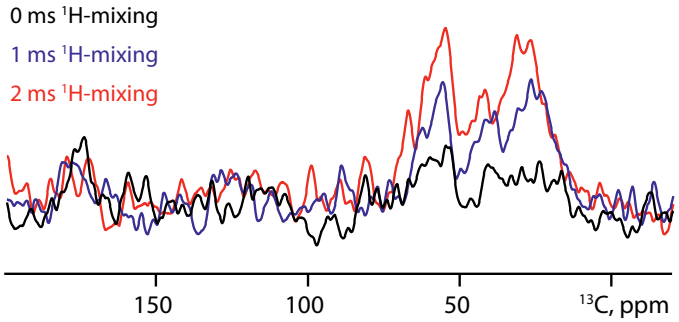
not present in the solution as the elution of the Ni-NTA column contains no MBP-Edc3 (lane 5). The second half of the incubated mixture of His-Edc3 and MBP-Edc3 homodimers was applied to an amylose column and a similar strategy was applied to probe for the occurrence of mixed dimers (lane 6,7). Taken together, we conclude that no exchange between Edc3 dimers takes place in 24 hours, indicating that the Edc3 dimers are very stable.



Supplementary Figure 2 | Comparison of 1D slices of NH spectra recorded by solution- or solid-state NMR. (**a-c**) 1D slices of resonances observed in both solution (green) and solid-state (gray) NMR spectra in Fig. 2a. Specific resonances are highlighted in Fig. 2a. Slices were scaled to match peak intensity.



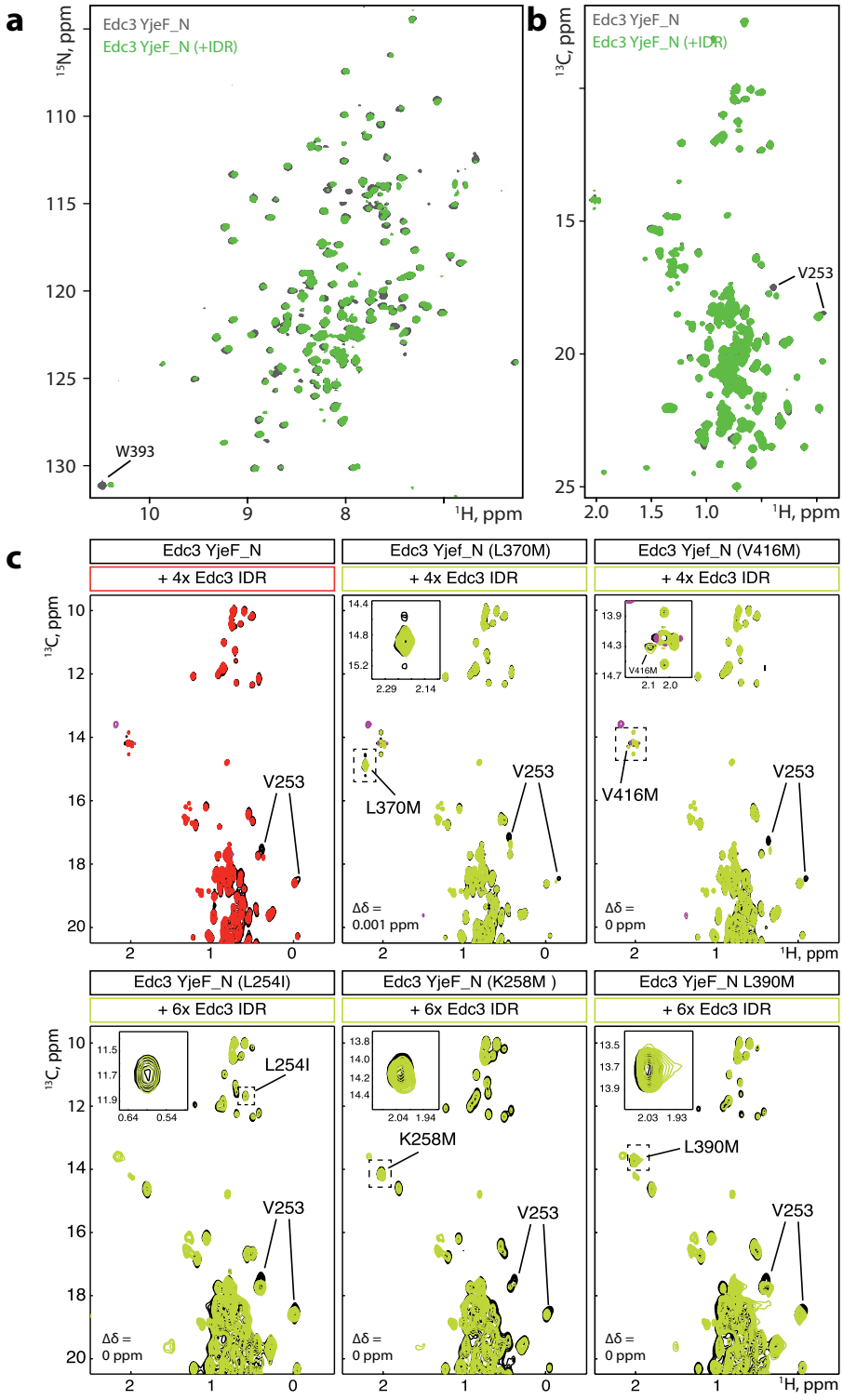
Supplementary Figure 3 | The fold of the YjeF_N domain in isolation is conserved in the matured state. Comparison of ^{1}H-detected dipolar NH (solid-state NMR) spectra recorded on Edc3-ΔLSm in the matured state (red) and lyophilized YjeF_N domain (cyan) highlighting the conservation of the fold of the YjeF_N domain. Differences between the two spectra can result from the IDR that interacts with the Yjef_N domain in the Edc3-ΔLSm construct as well as from differences in the ssNMR sample preparation.

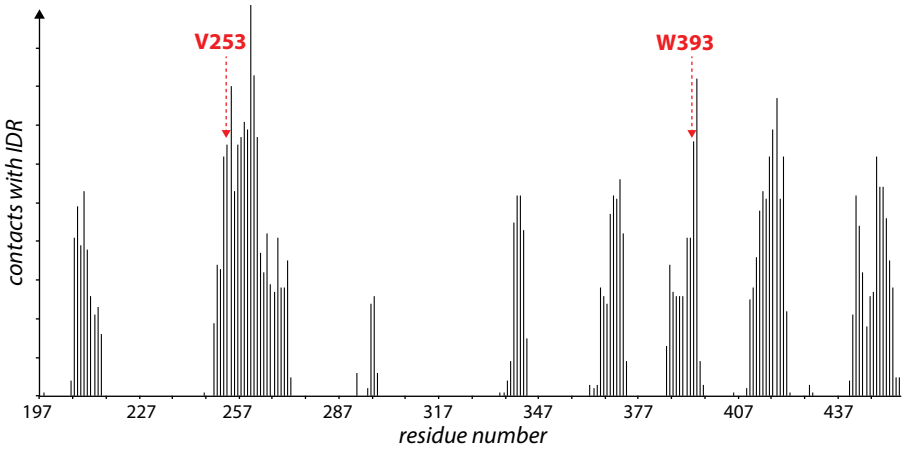


Supplementary Figure 4 | Edc3 dimers have intermolecular interactions during LLPS. NHHC spectra recorded on mixed ^{13}C-only labeled and ^{15}N-only labeled Edc3-ΔLSm with different proton mixing times shows the intermolecular interactions of Edc3-ΔLSm dimers in the matured state. Mixed ^{13}C-^{15}N Edc3 dimers are not present in the sample, as the dimers are highly stable (Supplementary Figure 1). The observed contacts are thus between protomers from two different Edc3 dimers. We have no indication that the Yjef_N dimers interact with each other, neither do we have an indication that two Edc3 IDRs interact with each other. Based on solution- state NMR data (Fig. 3a, 4a,b), we conclude that the Edc3 Yjef_N domain interacts with the Edc3 IDR. The inter-dimer contacts that we observe in the NHHC spectra are thus likely a result from interactions between the IDR of one Edc3 dimer with the Yjef_N domain of another Edc3 dimer.

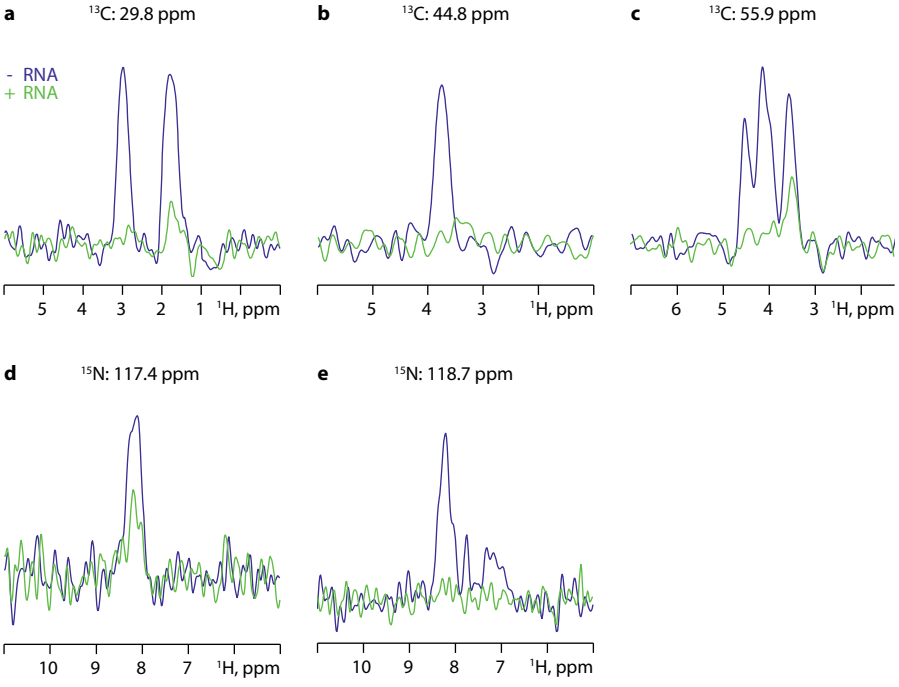
5

Supplementary Figure 5 | The Edc3 IDR interacts with the YjeF_N domain in solution. (**a**) Solution state ^{1}H-^{15}N TROSY and (**b**) ^{1}H-^{13}C HMQC spectra of the YjeF_N domain in solution in the absence (gray) or presence (green) of the unlabeled IDR. The assignment of W393 and V253 are highlighted as these residues show a significant chemical shift perturbation upon interaction with the IDR. (**c**) Introduced methyl groups on the Yjef_N surface that are not part of the IDR binding site (L370M, V416M, L254I, K258M, L390M). Note that the IDR interacts properly with the Yjef_N domain in all cases, as the naturally occurring methyl groups of Val 253 experience chemical shift perturbations comparable to the WT protein (panel b and Fig 4a). ➤


a
Edc3 YjeF_N
Edc3 YjeF_N (+IDR)
^{15}N, ppm
W393
^{1}H, ppm
b
Edc3 YjeF_N
Edc3 YjeF_N (+IDR)
^{13}C, ppm
V253
^{1}H, ppm
c
Edc3 YjeF_N
+ 4x Edc3 IDR
Edc3 Yjef_N (L370M)
+ 4x Edc3 IDR
Edc3 Yjef_N (V416M)
+ 4x Edc3 IDR
^{13}C, ppm
V253
L370M
V253
V416M
V253
Δδ = 0.001 ppm
Δδ = 0 ppm
^{1}H, ppm
Edc3 YjeF_N (L254I)
+ 6x Edc3 IDR
Edc3 YjeF_N (K258M)
+ 6x Edc3 IDR
Edc3 YjeF_N L390M
+ 6x Edc3 IDR
^{13}C, ppm
L254I
V253
K258M
V253
L390M
V253
Δδ = 0 ppm
Δδ = 0 ppm
Δδ = 0 ppm
^{1}H, ppm



Supplementary Figure 6 | Analysis of molecular dynamics simulations. The analysis of molecular dynamics simulations results in the identification of the binding epitope that was also determined based on solution-state NMR experiments (in red in Fig. 4c). The contacts between the IDR and the Yjef_N domain were summed over the last 50 ns of the MD simulations. Details of the MD results are provided can be found online.



Supplementary Figure 7 | RNA stabilizes dynamic unstructured domains of Edc3. (**a-c**) 1D slices of resonances with reduced intensity after addition of RNA (green) compared to the absence of RNA as observed in the scalar-based CH spectra (Fig. 5a) and (**d,e**) scalar-based NH spectra.

Supplementary Table 1 | Amino acid distribution and occurrence for the Edc3 YjeF_N domain and IDR.

	YjeF	*IDR*
Ala	26	7
Arg	6	3
Asn	14	9
Asp	10	9
Cys	3	1
Gln	16	3
Glu	8	8
Gly	18	6
His	7	4
Ile	18	2
Leu	33	8
Lys	10	20
Met	1	1
Phe	9	6
Pro	11	6
Ser	19	18
Thr	21	5
Trp	3	1
Tyr	3	2
Val	23	7

Supplementary Table 2 | Primer sequences of Edc3 cloning.

Primer	Sequence	Purpose
Edc3_NcoI_1_f	CATGCCATGGGTATGTCTGTAGCTGATTTTTATG	Cloning of Edc3 FL
Edc3_NcoI_72_f	GGCCATGGGTCCGAGTACAAATTCTACCAAGTTG	Cloning of Edc3 IDR and ΔLSm
Edc3_BamHI_194_r	CGGGATCCTCATTTCTTTAATGTTTCACGGCTG	Cloning of Edc3 IDR
Edc3_NcoI_195_f	GGCCATGGGCAAGCCATTCGTCGATGAAAA	Cloning of Edc3 YjeF
Edc3_BamHI_454_r	CGGGATCCTCAGTTAGTGCTGGTACATG	Cloning of Edc3 FL, YjeF and ΔLSm
IDR_del90-110_f	CGGAAACTTATTCTTCAAAGAACAAAACAAGTTTTCGCTGAATTTCG	Cloning of Edc3 ΔIDR1 mutant
IDR_del90-110_r	CGAAATTCAGCGAAAACTTGTTTTGTTCTTTGAAGAATAAGTTTCCG	
IDR_del158-171_f	CTGTTAAAGACGAATTTGTAGATTTAAGTTCTTCAAGCAATGGAC	Cloning of Edc3 ΔIDR2 mutant
IDR_del158-171_r	GTCCATTGCTTGAAGAACTTAAATCTACAAATTCGTCTTTAACAG	
YjeF_M-2G_fwd	CTTTATTTTCAGGGCGCCGGCGGCAAGCCATTCGTC	Binding site mapping
YjeF_M-2G_rev	GACGAATGGCTTGCCGCCGGCGCCCTGAAAATAAAG	
YjeF_V253I_fwd	CTTTCTCAATTTGTTTTTTCAATCTTAGGAGGACATAAACG	
YjeF_V253I_rev	CGTTTATGTCCTCCTAAGATTGAAAAAACAAATTGAGAAAG	
YjeF_L254I_fwd	CTCAATTTGTTTTTTCAGTCATTGGAGGACATAAACGATTGTC	
YjeF_L254I_rev	GACAATCGTTTATGTCCTCCAATGACTGAAAAAACAAATTGAG	
YjeF_K258M_fwd	CAGTCTTAGGAGGACATATGCGATTGTCTTCGAGG	
YjeF_K258M_rev	CCTCGAAGACAATCGCATATGTCCTCCTAAGACTG	
YjeF_L370M_fwd	GGCTAATGATTTAAACGTTATGATACTTTCGGTAGACATC	
YjeF_L370M_rev	GATGTCTACCGAAAGTATCATAACGTTTAAATCATTAGCC	
YjeF_L390M_fwd	CGGTACAAAAAAAAAACACGGCTATTATGCCCAAATGGACTTTGG	
YjeF_L390M_rev	CCAAAGTCCATTTGGGCATAATAGCCGTGTTTTTTTTTTGTACCG	
YjeF_K392M_fwd	CACGGCTATTTTGCCCATGTGGACTTTGGCTTTAGGG	
YjeF_K392M_rev	CCCTAAAGCCAAAGTCCACATGGGCAAAATAGCCGTG	
YjeF_W393M_fwd	GCTATTTTGCCCAAAATGACTTTGGCTTTAGG	
YjeF_W393M_rev	CCTAAAGCCAAAGTCATTTTGGGCAAAATAGC	
YjeF_V416M_fwd	GTTAAACAAGCTGCTGGTATGTCTGTATTTGTTGGAAAC	
YjeF_V416M_rev	GTTTCCAACAAATACAGACATACCAGCAGCTTGTTTAAC	
YjeF_L445I_fwd	CAAGTAACCGGCCAGTATATTGCCCAAATCTCATGTAC	
YjeF_L445I_rev	GTACATGAGATTTGGGCAATATACTGGCCGGTTACTTG	
YjeF_Q447I_fwd	GCCAGTATCTTGCCATCATCTCATGTACCAGCAC	
YjeF_Q447I_rev	GTGCTGGTACATGAGATGATGGCAAGATACTGGC	

Supplementary Table 3 | Protein constructs used in this study

Residue number / domain / mutations	Solubility / purification tag	Figure
1-454	N-His_6-TEV	1c, 1d, 1e, 3b, S1
72-194 (IDR)	N-His_6-GST-TEV	4a
195-454 (YjeF_N)	N-His_6-TEV	2a, 2b, 4a
YjeF_N (M-2G, L254I)	N-His_6-TEV	Supplementary Figure 5c
YjeF_N (M-2G, K258M)	N-His_6-TEV	Supplementary Figure 5c
YjeF_N (M-2G, L390M)	N-His_6-TEV	Supplementary Figure 5c
YjeF_N (M-2G, K392M)	N-His_6-TEV	4a
YjeF_N (M-2G, W393M)	N-His_6-TEV	4a
YjeF_N (M-2G, L445I)	N-His_6-TEV	4a
YjeF_N (M-2G, Q447I)	N-His_6-TEV	4a
YjeF_N (V253I)	N-His_6-TEV	4a
YjeF_N (L370M)	N-His_6-TEV	Supplementary Figure 5c
YjeF_N (V416M)	N-His_6-TEV	Supplementary Figure 5c
72-454 (IDR-YjeF_N)	N-His_6-TEV	1c, 1d, 2a, 2b, 5
72-454, Δ90-110 (IDRΔ1 – YjeF_N)	N-His_6-TEV	3b
72-454, Δ158-171 (IDRΔ2 – YjeF_N)	N-His_6-TEV	3b
1-454	N-MBP-TEV	Supplementary Figure 1
195-454 (Yjef_N)	N-MBP-TEV	Supplementary Figure 1

Supplementary Table 4 | RNA sequence used in this study

RNA	Properties	Sequence	Figures	Internal reference
30U15mer	GA-only, but single U at position 15	GGAGGAGAGGAAGGUAAGGGAAGAAAGAAG	5	#10 primer

Supplementary Table 5 | Overview of the solution-state NMR parameters

Experiment	Sample	Protein concentration	Scans	Acquisition Time	Window Function	Figure
NH TROSY	YjeF_N domain	100 µM	64	50 ms / 28 ms	QSINE (2)	Fig. 2a
Methyl TROSY	YjeF_N domain	100 µM	64	48.128 ms / 31.807 ms	QSINE (2)	Fig. 2b
Methyl TROSY	YjeF_N domain point mutants: V253I, L370M, V416M, W393M	100 µM	64	48.128 ms / 31.807 ms	QSINE (2)	Fig. 4 + Supplementary Figure 5
Methyl TROSY	YjeF_N domain point mutants: K259M, K392M, L254I, L390M, L445I, Q447I	50 µM	128	58.368 ms / 25.4453 ms	QSINE (2)	Fig. 4 + Supplementary Figure 5

Supplementary Table 6 | Overview of the solid-state NMR parameters

Experiment	Sample	Quantity	Scans	Acquisition Time (Direct-/indirect dimension)	Window Function	Figure
CC TOBSY	Edc3-FL	~ 5 mg	288	20 ms / 10 ms	QSINE (3)	Fig. 1c,d
CC TOBSY	Edc3-ΔLSm	~ 5.4 mg	288	20 ms / 10 ms	QSINE (2)	Fig. 1c,d
NH (Dipolar)	Edc3-ΔLSm Fraction Deuterated	~ 3.3 mg	480	30 ms / 15 ms	QSINE (6)	Fig. 2a
CH (Dipolar)	Edc3-ΔLSm Fraction Deuterated	~ 3.3 mg	352	19.7 ms / 7 ms	QSINE (3 and 6)	Fig. 2b and Fig. 5c-f
NH (Scalar)	Edc3-ΔLSm Fraction Deuterated	~ 3.3 mg	96	20 ms / 15 ms	QSINE (3)	Fig. 5b
CH (Scalar)	Edc3-ΔLSm Fraction Deuterated	~ 3.3 mg	96	20 ms / 6 ms	QSINE (3)	Fig. 5a
NH (Scalar)	Edc3-ΔLSm Fraction Deuterated + RNA	~ 3 mg	96	20 ms / 15 ms	QSINE (3)	Fig. 5b
CH (Scalar)	Edc3-ΔLSm Fraction Deuterated + RNA	~ 3 mg	96	20 ms / 6 ms	QSINE (3)	Fig. 5a
CH (Dipolar)	Edc3-ΔLSm Fraction Deuterated + RNA	~ 3 mg	480	21 ms / 6 ms	QSINE (4)	Fig. 5c-f
NH (Dipolar)	YjeF_N domain lyophilized	~ 4.8 mg	224	20 ms / 12.3 ms	QSINE (4.5)	Supplementary Figure 3



Chapter 6

Development of *in vitro*-grown spheroids as a tumor model system for ssNMR

Reiner Damman[1,2], Alessandra Lucini Paioni[1], Katerina T Xenaki[2], Paul MP van Bergen en Henegouwen[2], and Marc Baldus[1]

[1] NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht UniversityThe Netherlands.
[2] Cell Biology, Neurobiology and Biophysics, Department of Biology, Faculty of Science, Utrecht University, Utrecht, the Netherlands.

Manuscript in preparation

ABSTRACT

Recent advances in the field of in-cell NMR have made it possible to study isotope-labelled proteins in the context of bacterial or mammalian cell extracts or even entire cells. As most mammalian cells are part of a multi-cellular complex, there is a need to develop novel approaches to be able to study proteins at atomic resolution within the complexity of a 3D cellular environment. Here we investigate the use of the hanging drop method to grow spheroids which are homogenous in size and morphology as a model system to study solid tumors using NMR spectroscopy. We find that these spheroids are stable under solid-state NMR (ssNMR) magic-angle-spinning conditions for several days and show a clear change in metabolic profile as compared to single cell preparations. Finally, we utilize Dynamic Nuclear Polarization (DNP)-supported ssNMR measurements to show that low concentrations of labelled nanobodies can be detected inside the spheroids. Taken together, the development of in-cell ssNMR on spheroids opens a new field where proteins, or other biomolecules, can be directly studied in a 3D cellular environment with potential applications in pharmacological research.

INTRODUCTION

The contribution of cancer-derived cell lines to the understanding of molecular and cellular processes in the complex physiology of cancer has been immense. However, a full translation of the acquired knowledge towards patients can be challenging as the native tumor environment is much more complex than an in vitro monolayer culture. Due to the globular shape of a tumor, gradients exist for oxygen and nutrients, resulting in a different availability of nutrients for the proliferative outside layer as compared to the core[1]. Moreover, as a results of the Wartburg effect, synthesis of the endproduct of glycolysis, lactate, induces an acidification of the tumor and consequently results in altered gene expression and changes in the proteome. Furthermore, the extracellular matrix (ECM) forms a dense network of fibrils in which the tumor is embedded, driving tumor-specific intracellular signaling pathways[2–5].

The intricate interplay between the environment and the tumor plays a vital role in the survival of the tumor, although the consequences on a molecular level of protein functioning is poorly understood. In order to overcome these difficulties, a new dimension in cancer research has opened up by the advent of three-dimensional (3D) cell culture techniques. Several methods have been developed aiming at reproducing certain properties of the tumor and its microenvironment by *in vitro* cell culture, used in cancer research. Depending on the 3D cell culturing method followed, more complex aspects of cancer biology, like the effect of cellular architecture, cell-cell or cell-ECM contact or tumor microenvironment, can be studied[6–8]. Ultimately, patient-derived organoids are considered a highly complex system that offers the closest in vivo representation *in vitro*[9], making them optimal platforms to use in drug discovery. Overall, the microenvironment can have a prominent effect on protein function, supporting the necessity to study proteins within the most native context possible. Therefore, development of novel methods using 3D cell culture techniques for protein structural studies will contribute to a better understanding of the effect the microenvironment has on the structure-function relationship of proteins.

Recent advances in structural biology techniques allow for a rapid development of cellular structural biology. Due to breakthroughs in detection[10], image processing[11] and sample preparation[12], Cryo-Electron tomography (cryo-ET) has led to structural studies on the nuclear pore complex directly in mammalian cells[13,14]. However, due to the limited penetration of ~0.5 µm of the electron beam into the sample, cryo-ET is not sufficient to study large multi-cellular complexes. An alternative for cryo-ET is nuclear magnetic resonance (NMR), which is able to provide atomic level resolution without being limited by the sample size. Moreover, NMR requires isotope labeling of the protein of interest meaning no requirement for additional chemical modifications and finally, is non-invasive in nature. Tailored labelling approaches allowed for high-resolution spectroscopy bacterial membranes and could identify structural components at atomic resolution[15,16]. Selective labeling schemes of cytosolic proteins allowed for a characterization of thioredoxin and human FK506-binding protein in live cells[17]. Also for membrane proteins it has been shown that the native environment imposes changes to protein structure *in situ* and revealing new regions that that were not detected *in vitro*[18]. This notion is further supported by the increased line broadening of KcsA observed in bacterial membranes which suggests a structural heterogeneity[19,20] or increased dynamics of the non-activated extracellular domain of the epidermal growth factor receptor (EGFR) measured in cellular plasma membrane preparations[21]. Pioneering of delivery methods of spin-labelled material into unlabeled cells has resulted in the development of in-cell NMR for eukaryotic systems. A modified protein constructs provided

with a cell-penetrating peptide was able to translocate across the plasma membrane after which high-resolution spectra could be recorded[22]. In recent years, protein electroporation has been applied more frequently with much success, supporting the added value of studying protein structure and function with atomic resolution inside cells[21,23–26].

Although the significance of in-cell NMR is clear, research has focused on single cell preparations, yet the influence of 3D cell growth on proteins remains to be explored for proteins. However, solution state NMR has been used to study tumors and shown the necessity by highlighting clear cancer-related metabolic changes[27–29]. *In vitro*-grown spheroids also show similar changes in their metabolism, which supports the model systems as good mimetics of *in vivo* tumors and metabolic biomarkers can now be readily used to identify tumors[30,31]. A first report on the potential of atomic-resolution ssNMR-insights on proteins in tissue preparations, revealed strong correlations of murine ECM and tissue-culture derived ECM components[32,33].

Here we aim to develop a high-throughput approach to grown *in vitro* 3D tumor models which are homogenous in size and shape by adapting the existing hanging drop method[34]. Our data show that our spheroids remain stable during ssNMR measurements and show distinct metabolic differences as compared to monolayer cell cultures. Isotope-enrichment of amino acids in spheroids is readily achieved via the culture medium. Finally, addition of [^{13}C,^{15}N]-labeled anti-EGFR nanobody 7D12 penetrated into the spheroid[35], could be detected by sensitivity-enhanced NMR. Our data highlight the potential of 3D cell culture systems in structural biology and underlines the flexibility of studying labelled proteins in their native 3D environment. We believe that this will contribute to an increased understanding of the structure-function relationship of proteins.

RESULTS

A431 cells were grown in a monolayer culture until a confluency of 70-80% was reached. The cells were dissociated from the culture dish to be transferred into a hanging drop 384-well plate (Figure 1.A). Within 24 hours, several small spheroids were formed, which matured in the next 24-48 hours to a single spheroid. Sizes of spheroids can be controlled by the number of cells seeded in each individual well. To characterize A431 spheroids for ssNMR experiments, several spheroids matured spheroids were grown over a period of 48 hours. The spheroids were harvested from the hanging drop plates by centrifugation and collected for micro imaging by immobilizing them in a 1% agarose solution (Figure 1B). The spheroids were homogenous and did not differ in size or shape. To test whether spheroids can survive low speed magic angle spinning (MAS) for ssNMR experiments, spheroids were immobilized in a 1% agarose solution in a 3.2 mm rotor. Using micro imaging, images could be made of the spheroids before and after 5 kHz MAS at ambient temperature (Figure 1.C). The spheroids remained intact after MAS, although they were able to move through the agarose matrix during centrifugation. Confocal imaging of the spheroids, stained for nuclei and actin show tight networks of cells that retain the morphology of the spheroid during MAS (Figure 1.D). To test the intracellular changes, 1D HC INEPT spectra were recorded on unlabeled spheroids at 2 hr intervals to follow metabolic changes over time (Figure 1.E). No significant changes were observed during the recordings of the spectra, confirming that the spheroids remain viable. Taken together, our data show that in vitro grown spheroids using the hanging drop methods are homogenous and stable during ssNMR experiments.

For high-resolution and two-dimensional ssNMR experiments, incorporation of

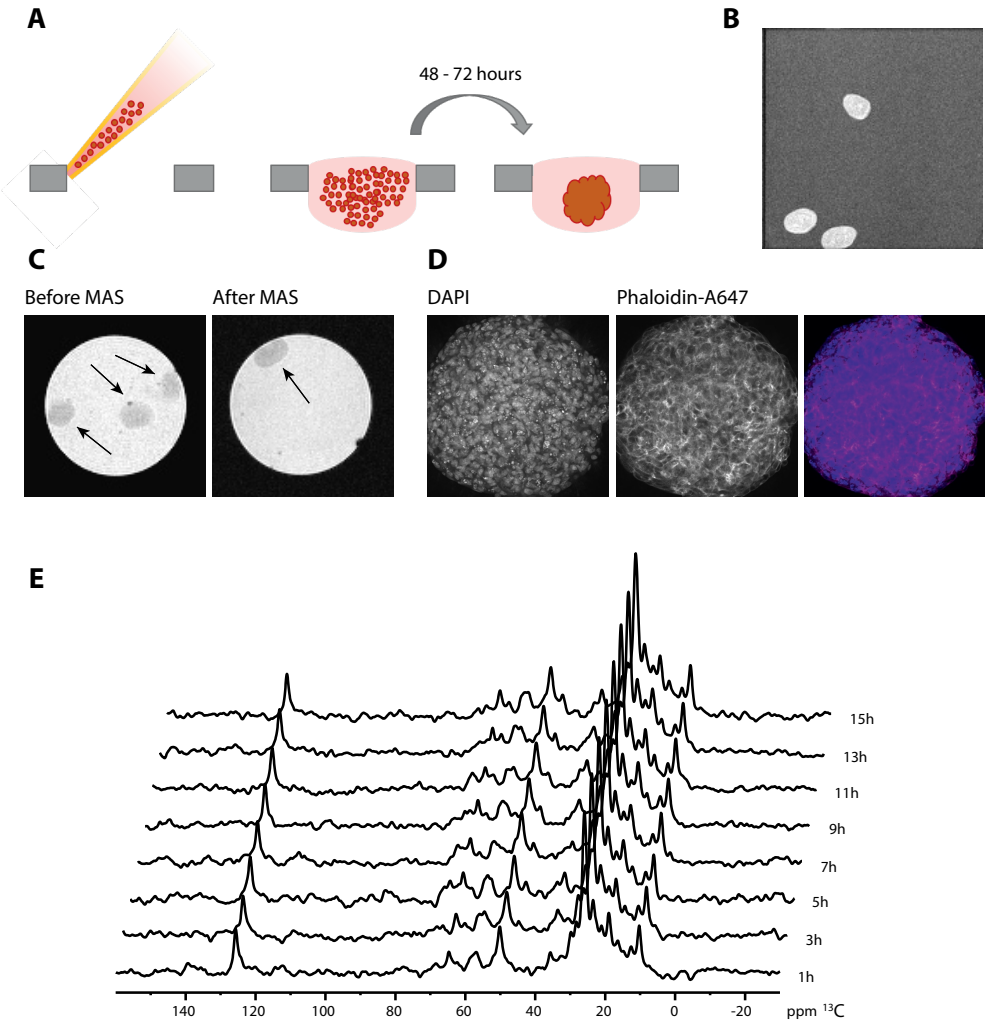


Figure 1 | *In vitro* grown A431 spheroids are stable under ssNMR conditions. A) A431 cells are removed from the monolayer culture dish and seeded in a well plate suited for hanging drop spheroid culturing. Homogeneous spheroids are formed between 48-72 hours after seeding of the cells in suspension. B) Side view of a micro-image of spheroids consisting of 100.000 cells per spheroid. C) A 3.2 mm rotor for ssNMR experiments was filled with spheroids consisting of 100.000 cells per spheroid. The spheroids were imaged from the top to show their morphology prior to 5.000 kHz MAS. Following MAS, the spheroids were reimaged to show that the morphology of the spheroids is maintained under MAS conditions. D) Confocal imaging of spheroids stained for nuclei (DAPI) and actin (phaloidin-A647) show the tight network of cells. E) Time course of 1D HC INEPT spectra recorded at a 2 hour interval.

[^{13}C,^{15}N] spin labels is required. A431 cells were grown in a monolayer culture for 2 passages on DMEM containing [^{13}C,^{15}N]-labelled amino acids, before they were seeded into hanging drop plates to form spheroids. Comparison of monolayer A431 cells and spheroids showed that the degree of labeling did not change upon the formation of spheroids and that the large majority of the rigidity-sensitive resonances originate from proteins (Figure 2.A,B). These data are in line with the nitrogen-detected cross polarization data showing clear backbone amide resonances and lysine and arginine side chain resonances (figure 2.C). Interestingly, 1D HC INEPT spectra show that there is a reduction in overall resonances observed in the spheroids sample as compared to the monolayer culture, suggesting that the spheroidal morphology affects the metabolites present (Figure 2.D).

To further investigate the effect of the spheroid culture on the metabolites inside the cell, two-dimensional HC INEPT spectra were recorded on monolayer- and spheroid cultures (Figure3.A). The majority of the resonances observed are present in both spectra. When comparing the ssNMR-spectra to a solution NMR HC HSQC spectrum of labelled growth medium, many of these peaks originate from the free amino acids in the growth medium. These data show that the isotopes in the growth medium are taken up by the cells and used by in cellular metabolism, mainly in protein synthesis. Interestingly, several peaks differ between the monolayer culture and the spheroid culture in either signal intensity of occurrence, where certain correlations are unique for their respective cell culture condition. For example, the spectral region up field of free amino acids shows 6 resonances that can be assigned to originate from a 6-ring sugar. These resonances are only present in the monolayer culture, while they are absent in the spheroid culture. The availability of nutrients is much reduced in the spheroid culture, which results in the total consumption of the sugar source. In conclusion, the labelled amino acids in the cell culture medium are consumed during proliferation and incorporated into several metabolic pathways, depending on their respective culture conditions.

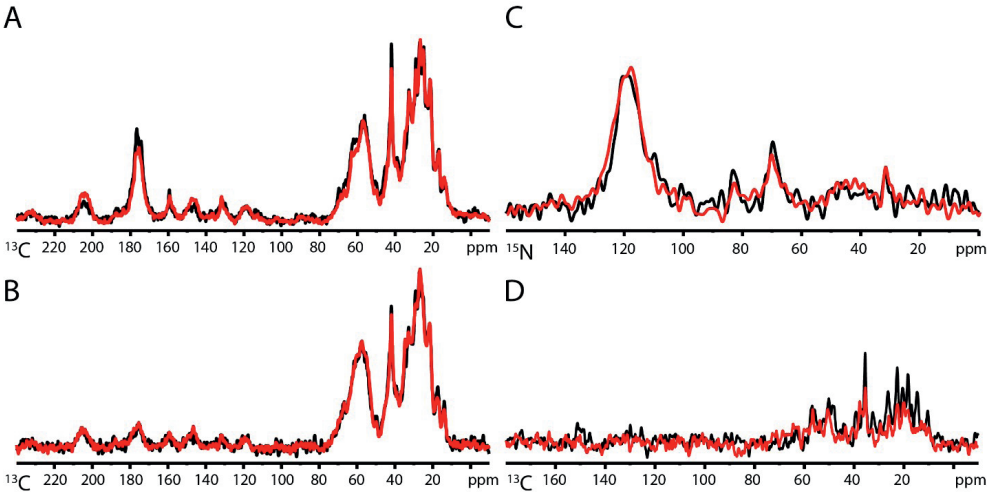


Figure 2 | Comparison of 1D solid state NMR spectra on fully labeled A431 cells. A-D) 1D ssNMR data recorded on fully labeled A431 cells grown in a monolayer (black) or in a 3D spheroid (red) showing ^{13}C 90 degree spectrum, ^{1}H^{13}C cross polarization, ^{1}H^{15}N cross polarization and ^{1}H^{13}C INEPT respectively.

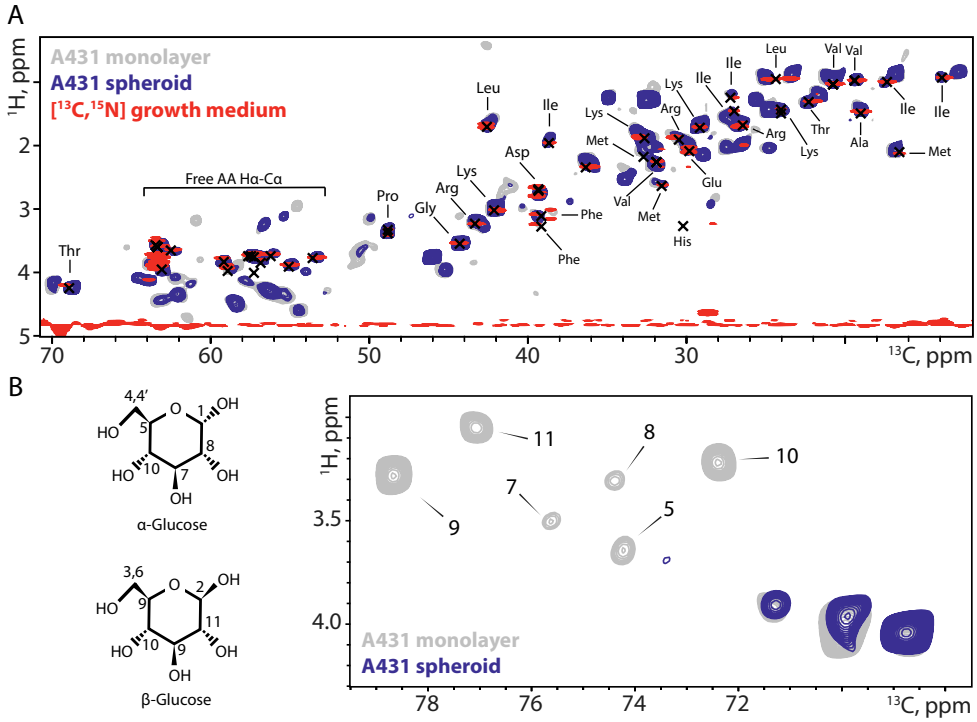


Figure 3 | A431 spheroids incorporate spin labels from the culture medium and show metabolic alterations compared to monolayer cells. A) HC INEPT spectrum recorded on labelled A431 cells from a monolayer culture and A431 spheroids at 5.000 kHz MAS. A CH HSQC of the culture medium (red) shows that the spin labels are incorporated into the monolayer (gray) and spheroid (blue) culture. The majority of the resonances match between the two culture methods, although several unique resonances are observed for the monolayer and spheroid culture, indicating differences in metabolism. Black crosses indicate solution state assignments as reported by the Human Metabolome Database. Note that glutamine, tryptophan and cysteine resonances are missing as they are added unlabeled. B) Zoom in on sugar region in the HC INEPT spectrum. The A431 monolayer culture shows several resonances which are consisted with HC correlation of 6-ring sugars. These resonances are absent in the spheroid culture.

As seen before, rigidity-sensitive carbon-detected spectra are dominated by protein resonances[21]. 2D CC PARIS spectra recorded on A431 spheroids, confirm the high degree of proteinaceous signals in the aliphatic region of the spectrum (Figure 4). For certain amino acid types, e.g. serine and threonine, (Cα-Cβ and Cβ,Cγ) cross peaks were observed that suggest different secondary structure elements were present which would be in line with the notion that the proteins inside the cells remain folded. Based on our previous experiments, actin is one of the most abundant protein inside A431 cells[21]. Hence, we used FANDAS[36,37] to predict chemical shift correlations using the structural model of actin[38]. The resulting correlation pattern is in qualitative agreement with the data recorded, suggesting that the CC PARIS[39] spectra are dominated by rigid proteins including actin. Due to the general labelling scheme, all proteins or other amino acid-derived metabolites are observed, making it difficult to study a single protein.

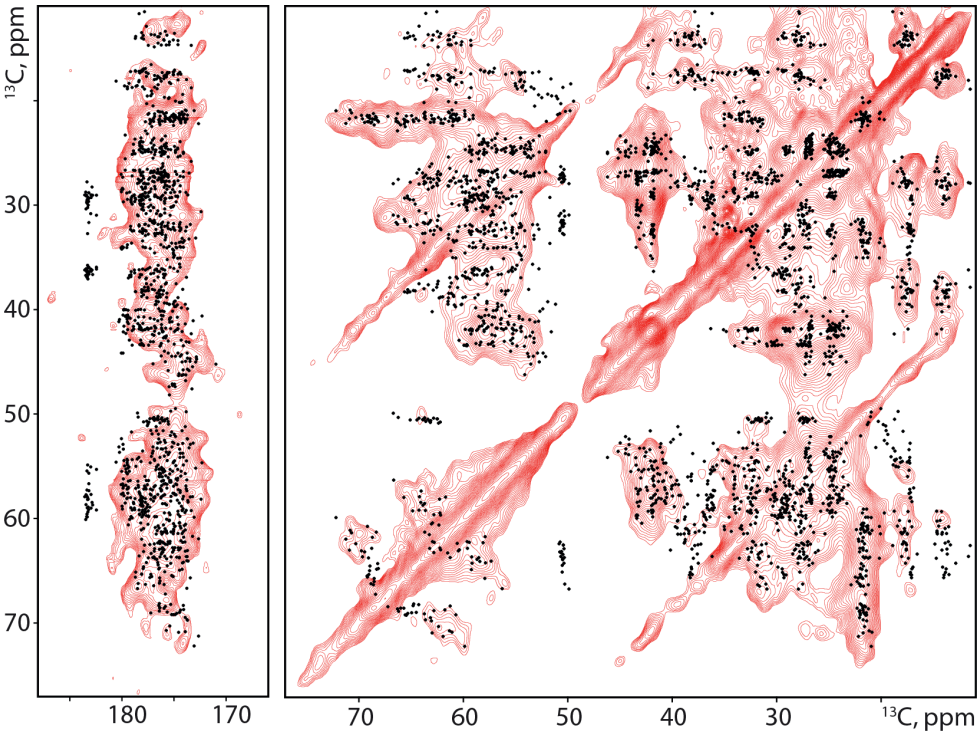


Figure 4 | A431 spheroids are enriched in labelled proteins which remain folded under MAS conditions. CC PARIS spectrum showing rigid protein signal of well-folded protein as seen by the dispersion of chemical shifts. Actin is the most abundant protein in cellular ssNMR samples from A431 cells. Therefore, chemical shift predictions of actin (black dots) are overlaid on top of the spectrum, based on the F-actin structure model (PDB ID: 2ZWH38). The observed resonances match well with the predictions.

In order to study single protein structures inside 3D cellular structures, we resorted to a different approach, where we added labelled protein to unlabeled spheroids. Previous work performed on nanobody 7D12 has proven that this is a small protein which binds with high affinity to its antigen EGFR[35,40,41], which is highly overexpressed in A431 cells. Spheroids were grown under normal culture conditions after which [^{13}C,^{15}N]-labelled 7D12 was added to the spheroids and allowed to penetrate into the spheroid for 1 hr. Unbound 7D12 was removed by replacing the culture medium with fresh culture medium without nanobody to allow the unbound nanobody fraction to diffuse out of the spheroid. We expect that the total amount of 7D12 inside the sample will be in the same order of magnitude as the amount of EGFR that is expressed in the spheroid, which will likely be too little for conventional carbon-detected ssNMR. In order to increase the sensitivity, we resorted to sensitivity-enhanced ssNMR, dynamic nuclear polarization (DNP), to enhance the overall signal. We observed that under DNP conditions, the overall protein enhancement observed is ~68 (Figure 5.A). Moreover, large enhancements were also observed for nitrogen-detected cross-polarization experiments where clear backbone amide and side chain nitrogen resonances are observed (Figure 5.B). To ensure that the NMR signal that we detected originates from 7D12 and not from natural abundance of the sample, we recorded 1D and 2D double quantum single

quantum (DQSQ) experiments using the SPC5 recoupling sequence[42]. These data show that we clearly observed protein signal and no other metabolites e.g. lipids or sugars, which suggests that we are able to detect bound 7D12 inside spheroids (Figure 5.C). Moreover, 2D DQSQ spectra reveal a close correlation with 7D12 chemical shift predictions based on the crystal structure of 7D12 bound to DIII of EGFR41 (PDB: 4KRL), confirming the presence of 7D12 (Figure 5.D). The additional signals observed around 70-78 ppm originate from natural abundance glycerol, which is added to the sample to prevent disruption of the spheroids. Taken together, our data show that small amounts of labelled material can be added to a complex environment and can be studied using sensitivity-enhanced ssNMR.

DISCUSSION

3D cell culture techniques, such as spheroids and organoids, are increasingly recognized as improved *in vitro* systems to study cells within their cellular context. This includes the effects of cellular contacts with other cells, extracellular matrix within tissues and specific growth conditions, such as low pH, present within the tumor micro-environment. The enhanced complexity contributes in the translation of *in vitro* data towards *in vivo* situations, as the heterogeneity in the sample will influence the structure-function relationship of proteins.

Currently, efforts are made to have a better understanding of the effect of the native environment on protein structure and functioning. However, it is often performed on bacterial- or mammalian-derived membranes for membrane proteins, or inside a single cell for cytoplasmic proteins. The additional complexity introduced by 3D cell culture techniques e.g. ECM, nutrient gradients, cell-cell contacts and changes in the metabolome could potentially have a strong influence on the structure-function relationship[43]. To obtain knowledge of these aspects, ssNMR can contribute by studying protein structure and dynamics at atomic resolution without the limitation of sample size. Here, we present a method where we grow simplified tumor model systems *in vitro* that are stable during the ssNMR measurements, which allows us to study protein structure in their most native environment.

Our results show that the hanging drop method for the generation of spheroids can be scaled-up to obtain large numbers of homogenous spheroids which can be readily manipulated and which are stable and remain intact during MAS for several days. Secondly, we observe clear differences in the metabolic profile of cells grown in a monolayer culture or 3D culture, emphasizing the necessity to explore 3D systems for structural biology. Thirdly, spin labels can easily be introduced into growing spheroids during culture and used to record 2D correlation spectra on proteins or other cellular components. Finally, labelled material can be added to unlabeled spheroids, which opens up new strategies to study a single protein of interest in the context of a complex environment.

The results presented in this study provide evidence that spheroids can be directly studied by ssNMR. Even though the spheroids presented here consist of the single cell type, the hanging drop method can be easily manipulated to include other factors that contribute to a tumor environment. Often, several types of cells are found within a tumor e.g. fibroblasts and immune cells, which can be included in the hanging drop method by adapting co-cultures protocols. Moreover, decellularizared ECM fibrils, such as collagen, could be added to the hanging drop, which will be incorporated by the growing spheroids to form a network of fibril- and cell-cell contacts. Additionally, advances in the delivery of labelled proteins into mammalian cells allow for new developments in sample preparation for sensitivity-enhanced ssNMR, so that individual proteins can be studied at atomic resolution. We envision that this

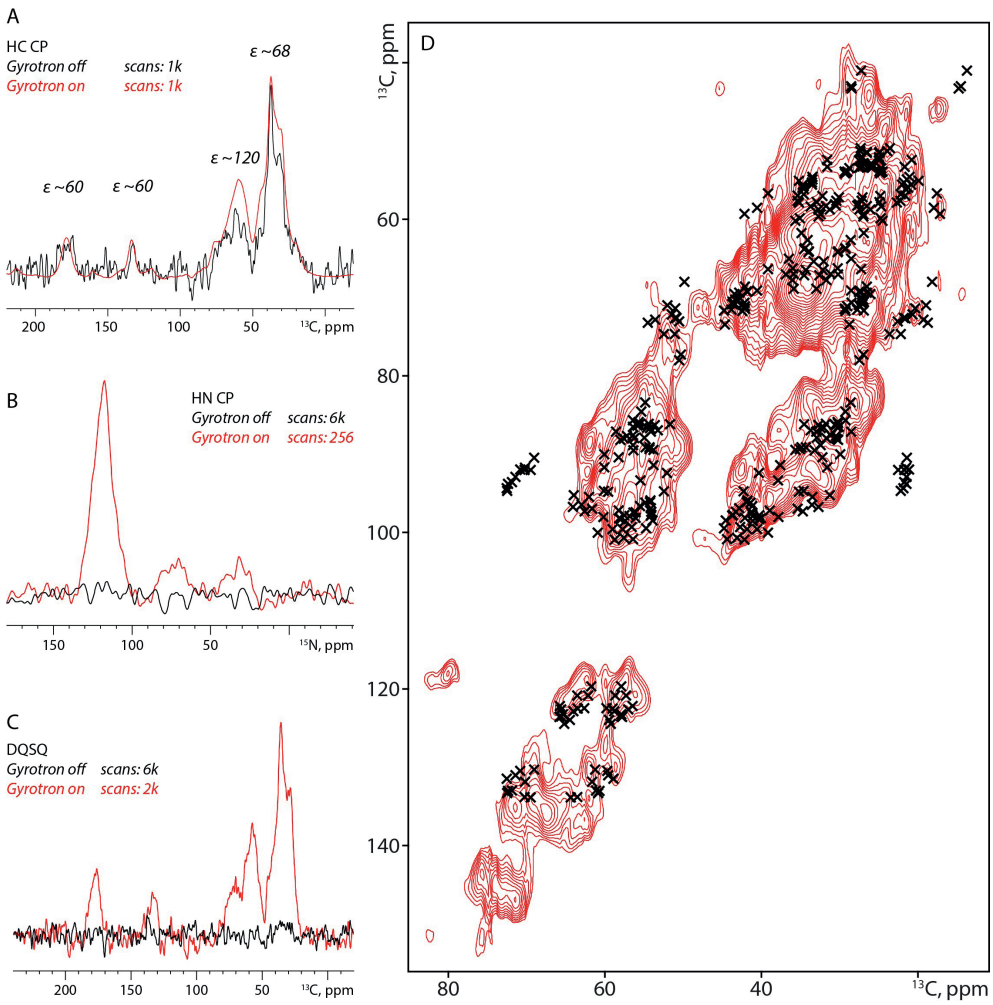


Figure 5 | DNP-supported ssNMR on labelled nanobody in spheroid. A,B) DNP-supported 1D HC (A) and HN (B) cross-polarization ssNMR experiments highlight the presence of labelled 7D12 inside A431 spheroids. The overall enhancement recorded on protein signal is 68 while on the glycerol is approximately 120. C) 1D DQSQ-filtered Carbon ssNMR spectra confirm labelled 7D12 inside the sample. D) 2D DQSQ spectrum recorded on 7D12-doped A431 spheroids highlighting the conserved fold of 7D12 inside spheroids. Black crosses represent chemical shift predictions of 7D12 based on its crystal structure, complexed to EGFR Domain III of the ECD (PDBidL 4KRL). All experiments were recorded on a 400MHz/GHz spectrometer at 100K.

will give rise to important insights into the role of the microenvironment on the structure-function relationship of proteins and contribute to pharmacological research where the gap in experimental systems between fields can be greatly reduced.

ACKNOWLEDGEMENTS

We thank Johan van der Zwan and Dr. Klaartje Houben for technical assistance. This work received financial support from the Dutch Science Foundation NWO (NWO-Groot grant 175.010.2009.002 and NWO TOP-PUNT grant 718.015.001 to M.B.) Experiments at the 950 MHz instrument were supported by uNMR-NL, an NWO-funded Roadmap NMR Facility (no. 184.032.207).

MATERIAL AND METHODS

Cell culture. A431 (ATCC, CRL-1555, LGC Standards, Germany) were grown in Delbecco's Modified Eagle's Medium (DMEM: GIBCO, Invitrogen, Paisley UK) supplemented 10% Fetal Calf Serum, L-Glutamine penicillin and streptomycin at 37°C with an atmosphere containing 5% CO_2.

Preparation of isotope-labeled spheroids. Detailed description of ^{13}C,^{15}N labeling of eukaryotic cells and medium preparation is described elsewhere (Kaplan, 2016[21]). In short, labeled culturing medium is prepared by dissolving 1 g/L of [^{13}C,^{15}N] algal amino acid mixture (Cortecnet) into DMEM absent of unlabeled amino acids. The medium is further supplemented with dialized FCS and penicillin and streptomycin. A431 cells were grown for 2-3 passages in labelling medium to enrich cells in ^{13}C,^{15}N-labelled proteins. For the generation of spheroids, A431 cells were grown to a confluency of 70% in a monolayer culture after which they were dissociated of the culture dish by incubation with Trypsin/EDTA and resuspended in fresh labelling medium. The cell suspension was seeded into 384-well Perfecta3D hanging drop plates (MERCK, HDP1385) and a concentration of 50.000 cells/well. The total volume of each well is 40 µL. Spheroids were grown for 48 to 76 hours at 37°C with an atmosphere containing 5% CO_2 after seeding. The spheroids were harvested by centrifugation for 5 minutes at 300x *g*.

Labeled nanobody production. Overnight cultures of *E. coli* BL21 codon+ cells contain plasmid DNA encoded for 7D12 behind a T7 inducible promoter were grown in LB containing kanamycin. The overnight culture was diluted (1/1000) into fully labeled M9 medium supplemented with 2 g/L ^{13}C-D-glucose and 0.5 g/L $^{15}NH_4Cl$. Bacteria were grown until an OD_{600} was reached after which protein production was induced by IPTG (0.5 mM) addition. 7D12 was produced overnight at 25°C. The next day, bacteria were harvested by centrifugation for 20 minutes at 4000x g and resuspended in ice-cold PBS. The cells were lysed by at least 2 freeze-thaw cycles followed by an additional centrifugation round of 30 minutes at 4000x g to harvest the periplasm. His-tagged 7D12 was purified from the periplasm by immobilized metal-affinity chromatography (IMAC). The HisTrap HP Ni^{2+}-Sepharose column was equilibrated with wash buffer (300 mM NaCl, 50 mM sodium phosphate, 10 mM imidazole, 2 mM β-mercaptoethanol). 7D12 was eluted from the column by the elution buffer (300 mM NaCl, 50 mM sodium phosphate, 100 mM imidazole, 2 mM β-mercaptoethanol). Protein fractions were pooled, and via buffer exchange, stored in PBS.

NMR and uMRI experiments. ssNMR and DNP experiments were conducted using 3.2 mm triple-resonance (^1H,^{13}C,^{15}N) magic-angle-spinning (MAS) probe heads at static magnetic fields ranging from 18.8 to 22.2 T corresponding to 700 MHz or 950 MHz proton resonance frequencies. Data were recorded at 273 K (referred as ambient temperature, AT) employing MAS rates of 5 kHz. Pulse schemes reflected standard homonuclear proton-driven spin diffusion (PDSD) and heteronuclear dipolar-based CH and NH cross-polarization (CP) sequences, and J-based CH sequence with INEPT transfer steps.

For µMRI experiments, spheroids were grown for 76 hours in normal DMEM after which they were harvested for sample preparation. The spheroids were washed once in ice-cold PBS and resuspended in Leibowitz 15 medium. The spheroids were immobilized in a solution of 2% agarose in Leibowitz medium inside of a 5 mm solution NMR tube. For the MAS test, spheroids were spun down in a 3.2 mm rotor and immobilized using 2% agarose in Leibowitz

medium. The 3.2 mm rotor was imaged inside of a 5 mm solution NMR tube.

Dynamic Nuclear Polarization. A431 spheroids were formed over a period of 76 hours after seeding in unlabeled DMEM. The spheroids were harvested by centrifugation for 5 minutes at 300 xg and washed warm (37°C) Leibowitz medium. [^{13}C,^{15}N]-labeled 7D12 (500 nM) was added to the spheroids in Leibowitz medium and allowed to penetrate and bind to the spheroid for an hour. The spheroids were washed again to remove the excess of nanobody. Then, unbound 7D12 was allowed to diffuse out of the spheroid during another incubation period in Leibowitz medium devoid of 7D12. Finally, the spheroids were washed again, pelleted and resuspended in DNP juice (60% deuterated glycerol in 10 mM HEPES, pH7.4 with 15 mM AMUPol).

DNP measurements were performed at 100K and 8 kHz MAS on a 3.2 mm triple-resonance (^{1}H,^{13}C,^{15}N) magic-angle-spinning (MAS) probe heads at static magnetic fields corresponding to proton/electron resonance frequencies of 400MHz/263GHz (Bruker BioSpin).

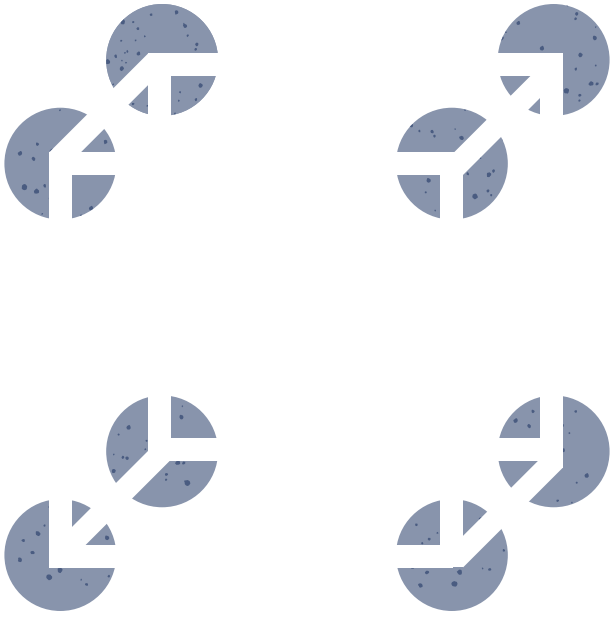
Microscopy Experiments. A431 spheroids were grown for 48-76 hours at a cell density of 4.000 cells/well. The spheroids were harvested and fixed for 30 minutes at room temperature using 4% paraformaldehyde solution. Nuclei and F-actin were stained using 1 µg/mL DAPI and 330 nM phalloidin-A647 for 5 minutes at room temperature. The spheroids were washed 3 times using ice-cold PBS and mounted with mowiol. Confocal images were obtained on a Zeiss LSM700 microscope.

REFERENCES

1. Kim, B. J. & Wu, M. Microfluidics for mammalian cell chemotaxis. Annals of Biomedical Engineering (2012). doi:10.1007/s10439-011-0489-9
2. Kim, W. et al. The integrin-coupled signaling adaptor p130Cas suppresses Smad3 function in transforming growth factor-α signaling. Mol. Biol. Cell (2008). doi:10.1091/mbc.E07-10-0991
3. Leight, J. L., Wozniak, M. A., Chen, S., Lynch, M. L. & Chen, C. S. Matrix rigidity regulates a switch between TGF-α1-induced apoptosis and epithelial-mesenchymal transition. Mol. Biol. Cell (2012). doi:10.1091/mbc.E11-06-0537
4. Liu, J. & Agarwal, S. Mechanical Signals Activate Vascular Endothelial Growth Factor Receptor-2 To Upregulate Endothelial Cell Proliferation during Inflammation. J. Immunol. (2010). doi:10.4049/jimmunol.0903660
5. Pylayeva, Y. et al. Ras- and PI3K-dependent breast tumorigenesis in mice and humans requires focal adhesion kinase signaling. J. Clin. Invest. (2009). doi:10.1172/JCI37160
6. Ravi, M., Ramesh, A. & Pattabhi, A. Contributions of 3D Cell Cultures for Cancer Research. Journal of Cellular Physiology (2017). doi:10.1002/jcp.25664
7. Rijal, G. & Li, W. 3D scaffolds in breast cancer research. Biomaterials (2016). doi:10.1016/j.biomaterials.2015.12.016
8. Weigelt, B., Ghajar, C. M. & Bissell, M. J. The need for complex 3D culture models to unravel novel pathways and identify accurate biomarkers in breast cancer. Advanced Drug Delivery Reviews (2014). doi:10.1016/j.addr.2014.01.001
9. Vlachogiannis, G. et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. Science (80-.). (2018). doi:10.1126/science.aao2774
10. McMullan, G., Faruqi, A. R. & Henderson, R. Direct Electron Detectors. in Methods in Enzymology (2016). doi:10.1016/bs.mie.2016.05.056
11. Ripstein, Z. A. & Rubinstein, J. L. Processing of Cryo-EM Movie Data. in Methods in Enzymology (2016). doi:10.1016/bs.mie.2016.04.009
12. Passmore, L. A. & Russo, C. J. Specimen Preparation for High-Resolution Cryo-EM. in Methods in Enzymology (2016). doi:10.1016/bs.mie.2016.04.011
13. Mahamid, J. et al. Visualizing the molecular sociology at the HeLa cell nuclear periphery. Science (80-.). (2016). doi:10.1126/science.aad8857
14. Maimon, T., Elad, N., Dahan, I. & Medalia, O. The human nuclear pore complex as revealed by cryo-electron tomography. Structure (2012). doi:10.1016/j.str.2012.03.025
15. Renault, M. et al. Cellular solid-state nuclear magnetic resonance spectroscopy. Proc. Natl. Acad. Sci. U. S. A. (2012). doi:10.1073/pnas.1116478109
16. Ward, M. E. et al. In situ structural studies of anabaena sensory rhodopsin in the E. coli membrane. Biophys. J. (2015). doi:10.1016/j.bpj.2015.02.018
17. Reckel, S., Lopez, J. J., Löhr, F., Glaubitz, C. & Dötsch, V. In-Cell Solid-State NMR as a Tool to Study Proteins in Large Complexes. ChemBioChem (2012). doi:10.1002/cbic.201100721
18. Kaplan, M. et al. Probing a cell-embedded megadalton protein complex by DNP-supported solid-state NMR. Nat. Methods (2015). doi:10.1038/nmeth.3406
19. Visscher, K. M. et al. Supramolecular Organization and Functional Implications of K + Channel Clusters in Membranes. Angew. Chemie - Int. Ed. (2017). doi:10.1002/anie.201705723
20. Medeiros-Silva, J. et al. 1H-Detected Solid-State NMR Studies of Water-Inaccessible Proteins In Vitro and In Situ. Angew. Chemie - Int. Ed. (2016). doi:10.1002/anie.201606594
21. Kaplan, M. et al. EGFR Dynamics Change during Activation in Native Membranes as Revealed by NMR. Cell (2016). doi:10.1016/j.cell.2016.10.038
22. Inomata, K. et al. High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. Nature (2009). doi:10.1038/nature07839
23. Binolfi, A. et al. Intracellular repair of oxidation-damaged α-synuclein fails to target C-terminal modification sites. Nat. Commun. (2016). doi:10.1038/ncomms10251
24. Frederick, K. K. et al. Sensitivity-Enhanced NMR Reveals Alterations in Protein Structure by Cellular Milieus. Cell (2015). doi:10.1016/j.cell.2015.09.024
25. Majumder, S. et al. Probing protein quinary interactions by in-cell nuclear magnetic resonance

spectroscopy. Biochemistry (2015). doi:10.1021/acs.biochem.5b00036

26. Theillet, F. X. et al. Structural disorder of monomeric α-synuclein persists in mammalian cells. Nature (2016). doi:10.1038/nature16531
27. Bollard, M. E. et al. Metabolic profiling of the effects of D-galactosamine in liver spheroids using 1H NMR and MAS-NMR spectroscopy. Chem. Res. Toxicol. (2002). doi:10.1021/tx025571e
28. Cox, B. L. et al. A novel bioreactor for combined magnetic resonance spectroscopy and optical imaging of metabolism in 3D cell cultures. Magn. Reson. Med. (2019). doi:10.1002/mrm.27644
29. Tee, S. S. et al. Targeted AKT inhibition in prostate cancer cells and spheroids reduces aerobic glycolysis and generation of hyperpolarized [1-13C] lactate. Mol. Cancer Res. (2018). doi:10.1158/1541-7786.MCR-17-0458
30. Ramachandran, G. K., Yong, W. P. & Yeow, C. H. Identification of gastric cancer biomarkers using 1H nuclear magnetic resonance spectrometry. PLoS One (2016). doi:10.1371/journal.pone.0162222
31. Ramachandran, G. K. & Yeow, C. H. Proton NMR characterization of intact primary and metastatic melanoma cells in 2D & 3D cultures. Biol. Res. (2017). doi:10.1186/s40659-017-0117-8
32. Chow, W. Y. et al. NMR spectroscopy of native and in vitro tissues implicates polyADP ribose in biomineralization. Science (80-.). (2014). doi:10.1126/science.1248167
33. Goldberga, I., Li, R. & Duer, M. J. Collagen Structure-Function Relationships from Solid-State NMR Spectroscopy. Acc. Chem. Res. (2018). doi:10.1021/acs.accounts.8b00092
34. Tung, Y. C. et al. High-throughput 3D spheroid culture and drug testing using a 384 hanging drop array. Analyst (2011). doi:10.1039/c0an00609b
35. Roovers, R. C. et al. A biparatopic anti-EGFR nanobody efficiently inhibits solid tumour growth. Int. J. Cancer (2011). doi:10.1002/ijc.26145
36. Narasimhan, S. et al. Rapid prediction of multi-dimensional NMR data sets using FANDAS. in Methods in Molecular Biology (2018). doi:10.1007/978-1-4939-7386-6_6
37. Gradmann, S. et al. Rapid prediction of multi-dimensional NMR data sets. J. Biomol. NMR (2012). doi:10.1007/s10858-012-9681-y
38. Oda, T., Iwasa, M., Aihara, T., Maéda, Y. & Narita, A. The nature of the globular- to fibrous-actin transition. Nature (2009). doi:10.1038/nature07685
39. Weingarth, M., Bodenhausen, G. & Tekely, P. Broadband magnetization transfer using moderate radio-frequency fields for NMR with very high static fields and spinning speeds. Chem. Phys. Lett. (2010). doi:10.1016/j.cplett.2010.01.072
40. Roovers, R. C. et al. Efficient inhibition of EGFR signalling and of tumour growth by antagonistic anti-EGFR Nanobodies. Cancer Immunol. Immunother. (2007). doi:10.1007/s00262-006-0180-4
41. Schmitz, K. R., Bagchi, A., Roovers, R. C., Van Bergen En Henegouwen, P. M. P. & Ferguson, K. M. Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. Structure (2013). doi:10.1016/j.str.2013.05.008
42. Hohwy, M., Rienstra, C. M., Jaroniec, C. P. & Griffin, R. G. Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy. J. Chem. Phys. (1999). doi:10.1063/1.478702
43. Siegal, G. & Selenko, P. Cells, drugs and NMR. J. Magn. Reson. (2019). doi:10.1016/j.jmr.2019.07.018



Chapter 7

General Discussion & Future Perspectives

Reinier Damman

NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands

GENERAL DISCUSSION & FUTURE PERSPECTIVES

Understanding the complex relationship between the structure of proteins and their respective function in cells is of great importance. The rigid structures that are obtained from classical structural biology techniques form a foundation to generate models of activation mechanisms of proteins, but generally lack the crucial native environment including complex lipid bilayer compositions in the case of membrane proteins and, in many cases, post-translational modifications. Previous studies have shown that the native environment can have major impact on protein structure, conformational dynamics and function. Therefore, the necessity to study the molecular activation in a native setting is considerable and will potentially lead to improved models on how proteins become are functioning[1,2].

The greatest challenge for performing high-resolution structural biology is the need for simplified *in vitro* systems for solution NMR, X-ray crystallography or (single particle) Cryo Electron Microscopy, which often requires additional stabilization of specific protein domains. Here, sold-state NMR (ssNMR) spectroscopy can offer an advantage over the other techniques as protein complexes can be studied independently of their dynamic behaviour via the combination of motion- and rigidity-sensitive experiments. Such a combination of ssNMR methods allows researchers to study and characterize their protein of interest across a wide range of motional timescales. Moreover, as ssNMR can in principle be used for a large range of molecular sizes, allowing the examination molecules embedded in larger complexes or associated with cellular membranes, or even whole cells[3,4]. In **Chapter 2** an overview is presented of recent advances in sample preparation and results for such *in situ*- and in cell-ssNMR on bacterial and mammalian samples. It illustrates how far the technique has come and highlights the challenges that need to be overcome in order to push the boundaries of in cell ssNMR further.

In this thesis, we set out to continue previous work from this lab in order to understand the molecular activation of growth factor receptors in their native environment and assess the importance of conformational sampling for receptor functioning. The role of dynamics in protein functioning was further assessed in a different system, where temporal sequestering of protein complexes on a mesoscale level has potentially large consequences on catalytic activity inside cells. Moreover, we developed novel approaches that pave the way for future in cell NMR studies.

The molecular activation mechanism of full-length EGFR has been characterized in a native lipid environment[1]. As the plasma membrane was directly harvested from mammalian cells, EGFR carries endogenous post-translational modifications, e.g. glycosylations. Interestingly, it has been shown that the EGFR ECD exhibits motion prior to ligand binding, which is lost after addition of EGF, showing that simultaneous binding of EGF to DI and DIII stabilizes the ECD. In **chapter 3**, we made use of allosteric nanobodies to reduce the conformational freedom of the EGFR ECD. Our data show that binding of EgA1 reduces conformational freedom of the ECD by locking the hinge region between DII and DIII. As EgA1 remains bound to the ECD, EGF can bind to the receptor as its binding epitope on DIII is still available for binding. Due to this lock, EGF cannot bind simultaneously to DI and DIII and subsequently, it is unable to induce EGFR functioning. This finding underscores the crucial importance of conformational sampling in EGFR activation, oligomerization and internalization. In line with this conclusion, EgA1 point mutants were created with the goal to abolish the interaction site on DII while leaving the DIII binding epitope unchanged. Interestingly, our data shows that binding of the EgA1 mutants increases the motion of the EGFR ECD, which

on its turn rescues the ability of EGF to activate EGFR again and to stimulate internalization. However, the binding affinity is reduced, suggesting that the ECD conformation that is stabilized by EGF differs from the crystal structure and most likely represents an asymmetric dimer as previously suggested for the EGFR dimer from Drosophila[5]. Even though the binding affinity is reduced after mutant EgA1 binding, the overall level of EGFR phosphorylation is increased, linking conformational freedom to EGFR functioning. This result is consistent with findings published by Orellana and colleagues who observed that activating point mutations found in glioblastoma patients show a larger structural diversity[6,7].

The freedom in conformational sampling of the EGFR ECD allows for stabilization of different ECD structures that trigger a diverse set of intracellular signalling cascades[8]. It has been proposed that dimer stability plays a role in the regulation of these signalling events and subsequent downregulation[7]. Our data supports this notion by modulating the conformational landscape by using allosteric nanobodies that directly affect the ECD motion. We propose that EGF can, besides a high-affinity structure, also stabilize a low-affinity structure leading to different outcome in EGFR oligomerization, activation and internalization. This might explain the increase in the co-called kiss and run of EGFRs in the extended conformation resulting in the observed increase in receptor phosphorylation and subsequent internalization. Moreover, the presence of different ECD conformers on the cell surface would allow for binding of a subset of ligands that potentially stabilize an ensemble of "open" ECD conformations, of which each structure potentially has unique characteristics, e.g. dimer stability, that might have different functional implications[8]. Taken together, the presence of conformational sampling of the ECD potentially allows for the formation of a diverse ensemble of conformations or conformers that directly affect EGFR functioning.

This opens the possibility that the amount of ECD dynamics correlates to EGFR functioning and therefore could offer an additional degree of regulation. Notably, EGFR function is influenced by specific glycosylation patterns and a local lipid environment that have the potential to alter EGFR ECD motion, thereby adding another level on control on receptor availability and activity[9]. The lipid environment might be controlled by the presence of specific lipids directly associated with the ECD such as GM1 or by the local environment provided by a lipid microdomain or raft such as sphingolipids in general and/or cholesterol. Hence molecular motion and receptor plasticity may expand the currently discussed structure-functioning relationship[10] and may providing further insights into aspects of receptor functioning such as binding affinity (including its thermodynamic description), conformational heterogeneity, oligomerization rates and states, dimer stability, phosphorylation, signalling and downregulation. Cells could be able to fine tune their local environment by changing post-translational modifications patterns or the lipid composition of the plasma membrane, thereby regulating receptor functioning, ranging from inhibition of activity to stimulation of receptor activation.

The crucial role of ECD dynamics of the molecular activation of EGFR also raises the question whether the conformational sampling is a conserved feature in growth factor receptors. To address this question, we set out in **chapter 4** to study another member of the EGFR receptor tyrosine kinase family. This receptor is Her2, which is expressed in 15-20% of all breast cancer patients. For these studies we used the breast cancer cell line SKBR3 which overexpresses Her2. From these cells, we isolated the Her2-containing plasma membrane fraction to study this receptor in the context of its native glycosylation- and lipid environment. Our ssNMR data confirm the presence of conformational sampling of the

unstimulated Her2 ECD. This is in contrast with the previous assumption that the ECD is stabilized in an "open" conformation[11]. As Her2 relies on heterodimerization for its phosphorylation, it is possible that ECD motion allows Her2 to adapt to different interaction partners within the ErbB family, but also interact with other receptors outside of the family, making Her2 heterodimerization and initiation of signalling a promiscuous event. The importance of ECD motion, or the change from a dynamic to rigid state, is further highlighted by the effect of adding the EGFR-specific allosteric nanobody EgA1. In **chapter 3**, we characterized the effect of the non-activating nanobody EgA1 on the EGFR dynamics and showed that the EGFR ECD motion is reduced after nanobody binding. The dynamics in the EGFR ECD could be rescued by introducing a point mutation in EgA1 that results in exclusive binding to DIII, thereby allowing for flexibility in the hinge between DII and DIII on EGFR. The data presented in **chapter 4** shows the indirect effect of EgA1 on Her2 functioning. Even though the allosteric inhibiting nanobody still binds to the EGFR ECD, we observe a clear phosphorylation of Her2, which could be abolished by the mutant EgA1. EgA1-induced Her2 phosphorylation was further supported by ssNMR data showing that the majority of the Her2 population becomes stabilized after incubation with EgA1. Taken together, non-activating ligands are able to induce phosphorylation Her2. We propose a mechanism where constitutively active Her2 is continuously probing rigid interaction partners. The conformational sampling of the Her2 ECD allows for rapid adaptation of its conformation to adjust its subdomains that lead to heterodimerization and phosphorylation. Taken together, the data presented in **chapter 3 & 4** underscore the fundamental requirement of ECD sampling in the regulation and control of receptor activation and signalling. As dynamics of the receptors may be affected by the native environment, the presence of the plasma membrane may be of crucial importance for a direct regulation of ECD dynamics.

Another demonstration of the importance of protein dynamics has been given in **chapter 5**. Here we present an atomistic model of liquid-liquid phase separation (LLPS) and propose a macromolecular structure adopted by Edc3 after LLPS. The combination of using solution- and solid-state NMR opened up the possibility to study and characterize large protein complexes, even when NMR assignments was largely missing. The direct comparison of proton-detected NMR spectra of truncated proteins in solution with protein complexes after phase separation (studied by ssNMR), paved the way to construct a model of how soluble proteins transit from solution to LLPS and to probe the role of each protein subdomain as well as the role of RNA in the process of LLPS. Moreover, we characterized several parameters that are of crucial importance for LLPS e.g. salt concentration or the hydrophobicity of the solvent.

In particular our combined approach such solution- and solid-state NMR data highlighted the vital role the intrinsically disordered region (IDR) plays in LLPS, a notion that was further confirmed by MD simulations. As interactions strengthen between the IDR and the folded dimerization domains, the overall dynamics inside the phase-separated state slow down until the point that this process is irreversible (gel-like state). Our current data suggest that this process does not involve major conformational changes of the IDR but more NMR experiments may be needed to fully proof this assumption. On the other hand, the N-terminal LSm domain of Edc3 is still mobile in the gel-like state, suggesting that it is still available to recruit and interact with Dcp2, the catalytically active protein. Moreover, LLPS relies on the accumulation of weak and likely unspecific interactions to form liquid droplets. Sections of the IDR remain mobile after LLPS which allows for the reversible character of LLPS *in vivo*.

The dynamic behaviour of protein subdomains after LLPS could also have significant implications in pharmacological studies, as LLPS has now been linked to neurodegenerative diseases, where protein aggregates and fibril deposits often give rise to the disease morphology[12]. When the aggregates are formed, solubilizing protein deposits has remained challenging. However, as the phase-separated state prior to fibril formation is dynamic and accessible for the diffusion of molecules, one could design novel molecular therapies that could tackle or prevent protein aggregation early on.

Finally, in **chapter 6** we describe a new approach which will allow researchers to perform NMR-based structural studies in *in vitro* grown tumour models. The spheroids form the basis upon which structural research can be performed. Even though we only have developed the first steps in the process, our results show the potential of using NMR on such a model system. The introduction of the 3rd dimension establishes two new parameters which could influence the structure and functioning of proteins: the introduction of gradients (e.g. nutrients, oxygen, pH, etc.) and the introduction of cell-cell contacts. Moreover, due to these new factors, cells that are in the outside layer exhibit a different metabolism and therefore also different proteome, lipidome and glycome in comparison to cells that are close to the necrotic core.

Labelling techniques still need to be developed which can provide structural information on a single protein species in a complex environment, for example a spheroid. Until then, one relies on the use of labelled molecules which can be added to spheroids. Pharmaceutical industries, as well as the societal mission to reduce animal model testing, there is an increased interest in studying 3D-cell culture model systems. The translation of *in vitro*-derived structural information is insufficient to correlate with these novel data. However, with the development of spheroids for structural studies, one could resort to these methods to study pharmaceutically relevant small molecules, peptides or proteins in the context of a 3D environment, where only the added molecules are labelled and therefore studied in an unlabelled cellular background. Alternatively, developing approaches to locally increase NMR sensitivity, by for example using nanobody-conjugates to bring a radical (DNP) close to the protein of interest, would contribute to study fully labelled cellular preparations. With these developments a single protein could potentially be studied in vitro but in conditions closely resembling physiological relevant states.

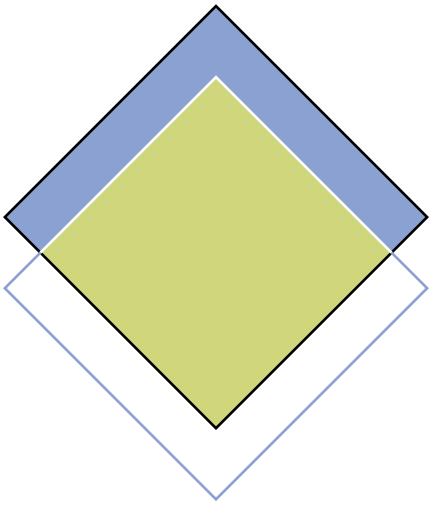
Chapters 3 and 4 show the crucial role that is played by the native environment on the functioning of growth factor receptors. The fact that EGFR and Her2 are often mutated in forms of solid tumours, e.g. head-and-neck cancer or breast cancer, presents the necessity to push the native boundaries even further. With the ability to study tumour model systems by ssNMR, we could envision to perform future studies on EGFR, Her2 or other growth factor receptors directly in 3D tumour models. In **chapter 6**, we used simplified systems in which a single cell line was used to grow the spheroids. This implies that the contacts cells make are only with other cells. Future research could continue to develop more complex tumour model systems by incorporating the extracellular matrix (ECM), which plays a major role in tumorigenesis. The ECM could be added in solution while the cells are seeded, allowing cells to incorporate ECM components into the spheroid. Alternatively, ECM-producing cell lines could be added to create a co-culture system where ECM components are produced and deposited into the extracellular space. When an intricate network of ECM is formed, intracellular signalling processes change which drive cell growth, sustained angiogenesis, evasion of apoptosis and inhibition of apoptosis[13]. Moreover, the physical barrier

imposed by the ECM combined with cell adhesion to the ECM drives resistance in tumours. Finally, there is direct evidence that EGFR and Her2 interacts with a subset of integrins, the molecules that connect the ECM and induce ECM-specific intercellular signalling events. Their interaction activates EGFR and induces cytoskeletal rearrangements involved in cell migration and recruits EGFR to focal adhesion hotspots. The change in EGFR localization potentially poses a new local environment that changes the EGFR ECD motion and directly also its functioning[14]. This effect has also been observed for Her2 in brain metastases[15].

Finally, in order to be able to study protein structure, dynamics and function in a native setting, ssNMR needs to develop more in order to be able to gain higher resolution data. First of all, currently it is a challenge to obtain a single labelled protein in the context of a cellular unlabelled background. Recent advances have highlighted the use of electroporation as a way to deliver labelled material into cells[4]. However, the labelled protein is still produce in bacteria, which means that it is reconstituted into a mammalian setting. Moreover, the life time of a protein is also too short to use this method for studies in spheroids as after the electroporation procedure, the spheroids still require 72 hours to form. Alternatively, one could resort to methods similar to those described in bacteria where after reach a certain OD in the bacterial growth curve, protein production is induced in fresh and labelled medium while endogenous expression is repressed, resulting in a high degree of labelling of your protein of interest[16]. Such systems still need to be established for mammalian systems, but offer the great advantage that the protein of interest is produced and studied inside the same cell. There is no need for purification of the protein. Also the ssNMR hardware is still under development with ultrafast MAS and proton-detection becoming more common, which give rise to higher-resolution data. Moreover, the recent development of the first ultra-high field 1.2 GHz NMR magnet scheduled for installation at Utrecht in 2021 will further expand the potential of NMR for *in-situ* structural biology in complex biomolecular systems.

REFERENCES

1. Kaplan, M. et al. EGFR Dynamics Change during Activation in Native Membranes as Revealed by NMR. Cell (2016). doi:10.1016/j.cell.2016.10.038
2. Staus, D. P. et al. Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. Nature (2016). doi:10.1038/nature18636
3. Medeiros-Silva, J. et al. High-resolution NMR studies of antibiotics in cellular membranes. Nat. Commun. (2018). doi:10.1038/s41467-018-06314-x
4. Narasimhan, S. et al. DNP-Supported Solid-State NMR Spectroscopy of Proteins Inside Mammalian Cells. Angew. Chemie Int. Ed. (2019). doi:10.1002/anie.201908849
5. Alvarado, D., Klein, D. E. & Lemmon, M. A. Structural Basis for negative cooperativity in growth factor binding to an EGF receptor. Cell (2010). doi:10.1016/j.cell.2010.07.015
6. Orellana, L. et al. Oncogenic mutations at the EGFR ectodomain structurally converge to remove a steric hindrance on a kinase-coupled cryptic epitope. Proc. Natl. Acad. Sci. U. S. A. (2019). doi:10.1073/pnas.1821442116
7. Bessman, N. J., Bagchi, A., Ferguson, K. M. & Lemmon, M. A. Complex Relationship between Ligand Binding and Dimerization in the Epidermal Growth Factor Receptor. Cell Rep. (2014). doi:10.1016/j.celrep.2014.10.010
8. Freed, D. M. et al. EGFR Ligands Differentially Stabilize Receptor Dimers to Specify Signaling Kinetics. Cell (2017). doi:10.1016/j.cell.2017.09.017
9. Coskun, Ü., Grzybek, M., Drechsel, D. & Simons, K. Regulation of human EGF receptor by lipids. Proc. Natl. Acad. Sci. U. S. A. (2011). doi:10.1073/pnas.1105666108
10. Kovacs, E., Zorn, J. A., Huang, Y., Barros, T. & Kuriyan, J. A Structural Perspective on the Regulation of the Epidermal Growth Factor Receptor. Annu. Rev. Biochem. (2015). doi:10.1146/annurev-biochem-060614-034402
11. Cho, H. S. et al. Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature (2003). doi:10.1038/nature01392
12. Alberti, S. & Dormann, D. Liquid–Liquid Phase Separation in Disease. Annu. Rev. Genet. (2019). doi:10.1146/annurev-genet-112618-043527
13. Pickup, M. W., Mouw, J. K. & Weaver, V. M. The extracellular matrix modulates the hallmarks of cancer. EMBO Rep. (2014). doi:10.15252/embr.201439246
14. Balanis, N. & Carlin, C. R. Mutual cross-talk between fibronectin integrins and the EGF receptor. Cell. Logist. (2012). doi:10.4161/cl.20112
15. Lal, S. et al. Interactions between αv-Integrin and HER2 and their role in the invasive phenotype of breast cancer cells in vitro and in rat brain. PLoS One (2015). doi:10.1371/journal.pone.0131842
16. Baker, L. A., Daniëls, M., Van Der Cruijsen, E. A. W., Folkers, G. E. & Baldus, M. Efficient cellular solid-state NMR of membrane proteins by targeted protein labeling. J. Biomol. NMR (2015). doi:10.1007/s10858-015-9936-5



&

Addendum

Summary
Samenvatting

Curriculum vitae
List of publications

Acknowledgments

NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands

SUMMARY

Cells rely on their ability to sense the environment in order to quickly and adequately respond to an ever changing environment. They are able to respond by means of proteins, which are the molecular machines of cells and allow for a proper cellular homeostasis. The relationship between the structure and function of proteins is complex but of crucial importance for cellular homeostasis. Activation of proteins frequently goes hand-in-hand with conformational changes as a consequence of ligand binding. Previous publications suggested that these conformational changes occur due to intrinsically dynamic protein domains. Often the physiological environment also imposes a significant effect on their functioning. When new proteins are synthesised, they are post-translationally modified in eukaryotic cells which can be important for their function. Moreover, membrane proteins are often embedded in specialized lipid environments that directly interact with the protein. Therefore, it is of great importance to study the structure of proteins under physiologically relevant conditions. This actually imposes a problem on structural biology techniques as they generally require a simplified and isolated system to obtain high-resolution data. In this sense, solid-state Nuclear Magnetic Resonance (ssNMR) spectroscopy is unique as it is not limited by the sample size and complexity. Additionally, ssNMR can provide insights into the dynamic- and rigid profile of proteins simultaneously. In this thesis, we describe studies where we studied the dynamics, activation and functioning proteins in a complex or native lipid environment using ssNMR and molecular biology techniques. These systems play a central role in the development of cancer, e.g. breast cancer, but are also involved in neurodegenerative diseases. In recent years there has been an increase in interest towards the development of novel ssNMR methods which opens up the possibility to measure cellular- and even in-cell-ssNMR samples.

In **chapter 2**, we present an overview of recent literature describing the development for prokaryote- and eukaryote samples. These methods offer the possibility to specifically label a single protein by enriching it with NMR-active isotopes inside a cellular system or they relate to approaches to deliver a NMR-visible protein inside an unlabelled cell. The literature overview emphasises the increased interest in this field and underscores the relevance of performing structural studies inside a native environment.

Previous work performed by Kaplan and colleagues highlighted the importance of the native environment on the intrinsic dynamic character of the extracellular domain (ECD) of the Epidermal Growth Factor Receptor (EGFR), but a role for this dynamics has never been described. **Chapter 3** presents new data, in which we made use of allosteric nanobodies showing that conformational freedom is essential in the functioning of EGFR, because the nanobody locks an important hinge region of the receptor. By making point mutations in the nanobody, we were able to increase the dynamic behaviour of the ECD which lead to increased activity and internalisation rates of EGFR. Because the dynamics of the ECD is crucial in the activation of EGFR, we wondered whether this is a general feature of growth factor receptors. In **chapter 4**, we adapted the method previously used for the EGFR studies, for a cell line which has an endogenous overexpression of HER2, a family member of the receptor tyrosine kinase family. HER2 is a unique receptor as until date, no activating ligand has been found. This implies that receptor activation is dependent on heterodimerisation. Here we present ssNMR measurements that, in contrast to the general assumption, suggest that the ECD of unstimulated HER2 is dynamic. EGF-dependent stimulation of HER2 results in an increase in phosphorylation levels and increased ECD rigidity. Finally, we made use of an anti-EGFR nanobody, described in **chapter 3**, which

induces a small rigidification of the EGFR ECD by locking a receptor hinge region. This nanobody was shown to indirectly increase phosphorylation levels of HER2 and also stabilize the HER2 ECD. We conclude that the presence of a rigid dockingpartner, and therefore not ligand binding, is the driving force for HER2 phosphorylation and oligomerization.

Chapter 5 describes a different system where the dynamic character is important in protein functioning. We made use of a combination of solution- and solid-state NMR to derive an atomistic model on phase separation. Our data shows that phase separation coincides with increased rigidity of a mobile unstructured region. We do not observed significant conformational changes. The intermolecular transient and weak interactions of the unstructured regions with the dimerization domain allow the formation of the dynamic macromolecular structure.Even though structural research is being performed in situ when possible, there is still a gap between the environment of cells within a human body and cells that are being cultured in a laboratory setting. Therefore, researchers in biology have developed 3-dimensional cell culture techniques which attempt to mimic the cellular environment inside a body. In **chapter 6**, we have developed a large-scale method which results in the formation of homogenous spheroids, which function as tumour model systems. We subjected them to ssNMR experiments in order to record structural data directly on the spheroids. Moderate MAS speeds did not disrupt the spheroid which enables structural studies directly on these model systems.

In conclusion, this thesis presents work which confirms the importance of the dynamic character of proteins and thereby links the structure and function of proteins. Additionally, it is important to perform such studies within the native environment. The combination of ssNMR and cell biological experiments was central in studying the structure-function relationship.

SAMENVATTING

Voor het overleven van cellen zijn zij erg afhankelijk van het waarnemen van- en reageren op hun omgeving. Dit doen zij vaak door middel van eiwitten, wat de moleculaire machines van de cel zijn. Om te kunnen begrijpen hoe deze eiwitten hun precieze functies kunnen uitvoeren, is het van belang om de structuur van het eiwit zo gedetailleerd mogelijk op te lossen. Echter, de de relatie tussen de structuur en functionaliteit van eiwitten is complex, maar desalniettemin van cruciaal belang voor een cellulaire balans. Wanneer een eiwit geactiveerd wordt door bijvoorbeeld ligand binding, leidt dit vaak tot conformationele veranderingen. Eerdere publicaties hebben gesuggereerd dat conformationele veranderingen mogelijk zijn omdat eiwitten intrinsieke dynamische domeinen hebben, waarna stablisiatie optreedt na ligand binding en daarmee het eiwit activeerd. Daarbij komt ook dat het fysiologische milieu van eiwitten vaak ook een grote invloed heeft op hun functionaliteit. Wanneer nieuwe eiwitten gesynthetiseerd worden, zullen deze in eukaryote cellen functionele additionele modificaties krijgen. Daarbij komt ook dat membraaneiwitten zich vaak in gespecialiseerde lipide-omgevingen bevinden waarbij de lipides een directe interactie hebben met het eiwit. Het is daarom van cruciaal belang om structurele studies uit te voeren onder fysiologisch relevante omstandigheden. Dit brengt echter een probleem met zich mee dat voor de meeste structerele biologische technieken gelimiteerd zijn door de specifieke eisen waarbij de monsters vaak vereenvoudigt worden en ze uit hun fysiologisch omgevingen worden gehaald. Vaste-stof Nucleaire Magnetische Resonantie (ssNMR) spectroscopie biedt echter unieke mogelijkheden voor structurele studies, omdat de complexiteit en formaat van een monster niet limiterend is. Daarnaast kan ssNMR simultaan informatie verschaffen over dynamische en rigide onderdelen van een eiwit. In dit proefschrift beschrijven we experimenten waarin we de dynamiek, activatie en functionaliteit bestuderen in complexe systemen middels ssNMR en moleculaire biologische proeven. Deze systemen spelen een centrale rol in de ontwikkeling van kanker, zoals borstkanker, maar ook in neurodegeneratieve ziektes.

In de afgelopen jaren wordt er meer energie geïnvesteerd in de ontwikkeling van nieuwe ssNMR methoden die het mogelijk maken om cellulaire of zelfs in-cel ssNMR monsters te maken. In **hoofdstuk 2** presenteren wij een overzicht van recente literatuur over ontwikkelingen in monsterpreparatie voor zowel prokaryote als eukaryote preparaten. Deze methoden bieden de mogelijkheid om bijvoorbeeld specifiek een enkel eiwit te maken in een cel dat verrijkt is met NMR-actieve isotopen of beschrijft de mogelijkheid om een NMR-zichtbaar eiwit in een cellulair systeem te krijgen. Het literatuuroverzicht benadrukt de toename in interesse in dit veld en de relevantie van dit type onderzoek voor structurele studies.

De fysiologisch omgeving is van cruciaal belang voor de intrinsieke dynamica van het extracellulaire domein (ECD) van de Epidermale Groeifactor Receptor (EGFR), maar een rol voor deze dynamica hiervan was nooit aangetoond. **Hoofdstuk 3** geeft nieuwe data weer waarin we, door middel van allosterische nanobodies, hebben kunnen aantonen dat de conformationele vrijheid van het ECD essentieel is voor het functioneren van EGFR. De allosterische nanobodies zetten een scharnier vast in het ECD waardoor EGFR niet meer functioneerd. Door het aanbrengen van puntmutaties in het nanobody konden we het dynamisch karakter van het ECD herstellen en zelfs versterken, wat leidde tot een toename van zowel de activiteit en internalisatiesnelheid van EGFR.

Doordat de dynamica van het ECD cruciaal is voor de activatie van EGFR, vroegen wij ons af of dit een algemeen eigenschap is van groeifactorreceptoren. In **hoofdstuk 4**

hebben we de methode die voorheen is gebruikt voor onze EGFR-studies toegepast op een nieuwe cellijn die een grote overexpressie heeft van HER2, een familielid van EGFR in de receptor tyrosine kinase-familie. De kristal structuur van HER2 ECD geeft een "open" conformatie weer, waardoor men denkt dat de ECD van HER2 rigide is en altijd klaar is om met andere receptoren een interactie aan te gaan. Daarnaast is HER2 een unieke receptor omdat er geen activerend ligand voor gevonden is, waardoor de activatie van de receptor afhangt van heterodimerisatie met andere receptoren. In dit hoofdstuk hebben we middels ssNMR-metingen aangetoond dat, in tegenstelling tot de algemene aanname, het ECD van HER2 ook dynamisch is voordat deze gestimuleerd is. Indirecte stimulatie van HER2 via EGFR door middel van EGF leidt tot een toename van het HER2 fosforylatieniveau en een toename in ECD-rigiditeit. Tot slot hebben we een nanobody gebruikt om de details van HER2 beter te begrijpen. In **hoofdstuk 3** hebben we dit remmende nanobody gekarakteriseerd en aangetoond dat de ECD van EGFR licht gestabiliseerd wordt. Het anti-EGFR nanobody was in staat om via indirecte stimulatie HER2 fosforylatie en ECD stabilisatie te induceren. We concluderen dat een rigide dimerisatie-epitoop, en niet ligandbinding, de drijfveer is voor het dimeriseren en functioneren van HER2.

Hoofdstuk 5 beschrijft een ander systeem waarin het dynamisch karakter belangrijk is voor het functioneren van het eiwit. We maakten gebruik van zowel vloeistof- als vaste-stof NMR om het proces van fasetransitie in detail te beschrijven. Onze data toont aan dat Edc3 in gist betrokken is bij het vormen van membraanloze compartimenten zonder grote structurele veranderingen. Echter, fasetransitie zorgt ervoor dat flexibele en ongestructereerde onderdelen van Edc3 tijdelijk betrokken zijn bij zwakke intermoleculaire interacties. De som van deze interacties leidt tot de formatie van een groter macromoleculair systeem die door het dynamische karakter nog steeds functioneel is.

Hoewel structureel onderzoek gedaan kan worden aan *in situ* monsters middels ssNMR, is er nog steeds een gat tussen cellen in het menselijk lichaam en cellen in het laboratorium. In de biologie wordt er daarom meer gebruik gemaakt van driedimensionele celkweeksystemen die potentieel een goede *in vitro*-representatie zijn van het lichaam. In **hoofdstuk 6** hebben we een methode ontwikkeld die het mogelijk maakt om op grote schaal homogene tumormodelsystemen te groeien en deze vervolgens te bestuderen middels ssNMR. Onze data bevestigen dat de 3D modelsystemen de experimentele condities van ssNMR overleven. Hierdoor is het mogelijk om bijvoorbeeld farmacologisch relevant onderzoek te doen, direct in een 3D tumor model systeem.

Samengevat presenteer ik in dit proefschrift onderzoek dat het dynamisch karakter van eiwitten cruciaal is: het verbindt de structuur en de functionaliteit van het eiwit. De fysiologische omgeving speelt hierbij een belangrijke rol en beïnvloed mogelijk direct de structuur en functionaliteit van eiwitten

CURRICULUM VITAE

Reinier Damman was born on the 4th of October, 1992 in Zwolle, the Netherlands. After completing his high school education in Thorbecke Scholengemeenschap in Zwolle, he pursued a Bachelor degree in Biology at the Rijksuniversiteit Groningen, with a specialisation in molecular lifesciences. For his Bachelor internship, he attempted to reconstitute proteins inside nanodiscs under the supervision of Dr. Jeanine de Keyzer and Prof. Dr. Arnold Driessen. After he finished his Bachelor's in 2013, he went to Utrecht University where he continued to pursue his interest in molecular biology in the Master Molecular and Cellular Life Sciences. During his first Master internship in the lab of Dr. Paul van Bergen en Henegouwen, he set out to understand the molecular determinants in Epidermal Growth Factor Receptor (EGFR) clustering. He was involved in the establishment of quantitative fluorescent imaging techniques which report on oligomer size via FRET between the same fluorophores. Besides, he also got involved in optimizing a quantitative imaging technique that counts GFP molecules inside a cluster to determine oligomer sizes. Afterwards, he went for his second Master internship to Genmab, a pharmaceutical antibody company, to set up high-throughput screening methods for combination therapy using small molecules and antibodies.

Reinier received his Master degree in 2015 after which he joined the lab of Prof. Dr. Marc Baldus at the NMR Spectroscopy department, Bijvoet Center for Biomolecular Research at the Utrecht University, the Netherlands, where he became a shared PhD student with the lab of Dr. Paul van Bergen en Henegouwen at the Cell Biology division, Department of Biology Faculty of Science, Utrecht University, The Netherlands. His main goal was to unravel the complex relationship between the structure and function of growth factor receptors, which drives the molecular activation of these proteins. He was also involved in research that led to the development of a structural model on protein structure after phase separation. Reinier also developed novel methods which allow future researchers to perform structural studies directly on three-dimensional tumour model systems. He presented his work at various national and international conferences and was awarded a poster prize at the Science4Life conference in 2018, and was involved in the supervision of Bachelor- and Master students. The results of his work are described in this thesis.

LIST OF PUBLICATIONS

Allosteric nanobodies reveal the importance of regulating conformational dynamics in EGFR functioning
Damman, R., Kendek, A. A., van Bergen en Henegouwen, P. M. P., & Baldus, M.
In preperation

Development of *in vitro*-grown spheroids as a tumor model system for ssNMR
Damman, R., Lucini Paioni, A., Xenaki, K. T., van Bergen en Henegouwen, P. M. P., & Baldus, M.
In preperation

Chapter 9 Cellular Solid-state NMR Spectroscopy. in In-cell NMR Spectroscopy: From Molecular Sciences to Cell Biology
Damman, R., Narasimhan, S., Weingarth. M., & Baldus, M.
The Royal Society of Chemistry (2020) 131–151

Atomic-level insight into mRNA processing bodies by combining solid and solution-state NMR spectroscopy.
Damman, R., Schütz, S., Luo, Y., Weingarth, M., Sprangers, R., & Baldus, M.
Nature Communications. (2019) 10, 4536

EGFR Dynamics Change during Activation in Native Membranes as Revealed by NMR.
Kaplan, M., Narasimhan, S., de Heus, C., Mance, D., van Doorn, S., Houben, K., Popov-Čeleketić, D., Damman, R., Katrukha, E. A., Jain, P., Geerts, W. J. C., Heck, A. J. R., Folkers, G. E., Kapitein, L. C., Lemeer, S., van Bergen en Henegouwen, P. M. P., & Baldus, M.
Cell (2016) 167(5):1241-1251

ACKNOWLEDGMENTS

And then the end is suddenly here. It feels weird writing this part as it is a direct confrontation with reality that the PhD will be finished soon. For a long time, this moment has been something that was always far away in the future. It has been quite the journey and not always a smooth ride, but I've come to believe that this is inherent to doing a PhD. As a freshy graduated masters student I jumped into the deep-end. As Mr. Baggins rightfully said "I'm going on an adventure!". Maybe a bit naïve, I thought the acquisition of knowledge would be central during the period. Looking back at it though, I've come to the realization that there is much more to it and I've learned more about myself then I probably even realize at this moment. Luckily, I was fortunate enough to have many friends, family and colleagues around that could prepare me well prior to this adventure and could help me as well along the way. I would now wish to take a moment to thank all of you.

First off all, I would like to thank the two people that made it all possible, my promotor Marc and my co-promotor Paul. **Marc**, thank you for allowing me the opportunity for doing my PhD in your lab. I can imagine that on a first glance, I'm not the ideal candidate. Only a few encounters with NMR prior to the PhD, which makes my starting point on the already steep learning process, even worse. Luckily, you gave me the time to acquire the necessary skills for recording the required data. It is very likely that I'm close the worst spectroscopist that walked through these halls, but I do hope that the work I did is a foundation where future candidates can benefit from. You allowed me to develop myself as a scientist in a field that is very complex and multidisciplinary, making it not always easy to have all the pieces of the puzzle together. Yet our discussions were fruitful leading to new questions to chase in the lab. I have always been impressed by the diverse range of research your performing, trying to tackle all of these open questions, besides all the other responsibilities you have. I would like to thank you for these 4+ years and wish you all the best for your future.

Paul, I remember the first time we met still very well. I just started my Master's in Utrecht and was enrolled into the first course of the year. There you presented work from your lab about EGFR, where I was intrigued by the beauty and complexity of this receptor. Especially, the discussion about the molecular mechanism of receptor-induced internalization fascinated me. This was the moment where EGFR grabbed and apparently never let go as I'm still working on it. Thank you for keeping me in mind for this position. I worked for years happily on these projects, getting lost at times, drowning in details about molecular behavior. You gave me the freedom and trust to perform these experiments, although sometimes skeptical, and for our scientific discussions. What I always liked is your unconventional thinking towards projects, research and science. It is good to hear a different point of view to force you to get outside of your comfort zone. I very much appreciated working together with you during these years and wish you a bright future.

Next, I would like to thank the group leaders from the NMR department & Cell Biology departments with whom I had the pleasure to work the most with. **Markus**, you were the one that had the patience to show me the ropes of recording the NMR experiments. Though times I couldn't follow you completely, you managed to teach me the skills required to record spectra. Moreover, I happily worked together with you on the Edc3 project and would like to thank you for all your efforts. Good luck in the future, I'm sure the Weingarth Lab will do well. **Alexandre**, the captain of the computational ship. Your research is fascinating, especially the incorporation of biochemical data into the docking. Besides, I am amazed by how you manage to travel the world while always having time for your lab. Thank

you for being part of my committee and I wish you all the best. **Hugo**, halfway though my PhD you made the transition from Leiden to Utrecht, a great addition to the NMR team. I admire your broad knowledge ranging from biochemistry in the wet lab to NMR. Moreover, your scripts made solution NMR very accessible that even I could do it. I wish you all the best for your future. **Sabrina**, even though some of the projects you do share the same receptor as what I studied, our research differs tremendously. It is inspiring to see you trying to develop novel therapeutics using nanobodies and hope that you'll be successful in doing so.

To **Remco Sprangers**, **Rolf Boelens**, **Alaxandre Bonvin**, **Piet Gros** & **Boudewijn Burgering**, thank you for being part of the reading committee and I am looking forward to discussing the work with you in the near future. **Remco**, I would also like to thank you for the fruitful collaboration we had. Phase separation is a fascinating concept of which we only understand so much. **Rolf**, your wisdom of and passion for NMR is contagious and I would like to thank you for your private lectures to help me understand it better.

All of the work I did in the two labs could not have been performed without the expertise and commitment of the following group of people. **Johan**, you are the silent force of the NMR group. Your knowledge about every piece of hardware that is necessary for recording spectra is immense, from sample prep to maintenance to making new coils, you do it all. Thank you for all your help during the times that I was panicking when things wouldn't work in my hands. **Mark** (you bloody !), it has been a pleasure to work alongside with you, even though we hardly worked directly together on any project, besides the kilo of nanobodies you produced for the lab. Echt "Hééél veel" nanobodies. The lunches have been legendary and left me often with a stomach pain from laughing. Thank you for your support during a wandelingetje. It hasn't been the same since you left but I'm happy you found a new challenge. **Klaartje**, you were here for the first period in my PhD. Your incredible diverse background and understanding of both NMR and biology made it easy for me to get used to this new environment. Thank you for the discussions and input and getting me up to speed. **Gert**, you're a millipede in the lab. Supervising students in the lab while performing your own experiments in the lab and teaching at the same time. It's impressive you get it all done. Also your unique angle at research challenged me to see experiments in a different light. **Hans**, thank you for the fun times together at borrels and other social events. Your comprehension of research and down-to-earth personality made it a pleasure to work alongside each other. **Andrei**, I would like to thank you for all the work that you've put in since you arrived and the information you shared about hedgehogs. I won't forget that anymore. **Raymond**, you only just started so we didn't have much time yet to interact. I wish you all the best in the department and good luck in the future. **Sófia**, thank you for helping me out in the lab and not batting an eye everything I came begging for nanobodies again. You bring the lab to live when you're singing along to the Red Hot Chili Peppers or deciding on the next "Crush of the Month". I wish you all the best on your new adventure! **Barbara**, thank you for the lovely times we had during the social events and everything that you did for me in the beginning of my PhD. **Geeske**, you are Fryslân to the bone. Your car even originates from there! It has been a pleasure to get to know you and enjoy our morning coffees. I wish you, Frans and the boys all the best for the rest of your lives.

My journey started on the same day is my office mate **João**, although for a lot of people more frequently known by many other names. From the beginning you established yourself as a central player within the NMR groups. You're kind and always make time in your

busy schedule. I say busy, because man you work. Luckily at the end of your PhD it also shows with many nice papers. You're commitment for science is admirable but that doesn't come at the cost for being there for others. I often had a (dumb) question for you which you gladly answered. You're now going to start a new adventure far away, all the best with that! All that I want to say still is don't forget to live and be a good boy!

I had the pleasure of sharing my PhD with some of the most amazing people. On the same day that I started, also **Yanzhang** (Jonny) begin his adventure. We knew each other since our Master's as you also did an internship in the lab of Paul together with Rachid. Jon you were a silent force to be reckoned with. I think you managed to work very efficiently, not complaining much along the way, yet producing well. In the beginning you were maybe a bit shy, but that changed completely during your time here. I cannot count the amount of times we laughed so incredibly hard about the inside jokes. You're hilarious. I wish you and Jane all the best, and one last time "NGK!". Together with Jon, Marc referred to the three of us sometimes as the three musketeers. **Siddarth** (Sid), you started a little bit earlier than me with your PhD. After 2 successful internships in Bloembergen you decided to stay but change your research completely. The projects were very challenging to setup up but it is impressive to see how far you managed to push things. This is also recognized by the community. You should be proud of what you're doing! You are a great guy who will always be there for you. I could discuss my struggles with you frequently and you managed to motivate me again. Good luck with all the things you're going to do! **Alessandra** (Ale), always happy, always cheerful. I cannot remember a day where I did not see you smiling. I enjoyed working together with you in some challenging projects. Thank you for the lovely time and good luck with everything you do. **Barend** (Bährund), thank you for the time we worked together. You're humor is off the charts and together with Mark D. unstoppable. These moments are missed! **Helena** (Leona), I will never forget your laugh as you are precessing like a spin in a magnet. Also, the time you smacked me in the face, hard to forget. Your handprint still burns when Düsseldorf comes up. **David**, the force from Canada. Good luck for you PhD and I'm sure you'll do fine. **Agnes** (Agata), you are a great addition to the lab. Your personality combined with the lab skills will be appreciated. All the best with separating phases. **Anamika**, you're the one to whom I'll hand over the metaphorical torch. You've impressed me with how fast you pick up things and your level of organization. Your own label maker, WHAT?! I wish you all the best for your PhD and keep smiling.

Unfortunately, as the PhD is a (relatively) fixed amount of time, people will also leave the lab. **Mohammed**, you showed me the ropes of the EGFR project. You got me up to speed quickly with the protocols you set up which allowed me to continue smoothly. **Deni,** your NMR skills were insane. The comprehension of these concepts was impressive, but to be able to explain that to NMR newbies in language that one can understand is admirable. Thank you for taking the time helping me settling in and also for working together on some projects. **Cecilia**, (Pinto!) as the other Portuguese person in the lab I had the pleasure to every now and then practice some of my (lacking) Portuguese skills. Queen of the Bam sample prep, I remember the amount of different samples you prepared using complex and long protocols. Since you left the hallways are much cleaner (no more coffee?!) and I still can't believe you don't like bitoque. **Maeghan**, thank you for the addition of your expertise in cellular NMR and happy personality. The freezer is still working! **Shengqi**, a hardcore spectroscopist. Thank you for the discussions and good luck to you with your own lab.

Also my other lab family helped me a lot during my time. **Katerina,** you are a

prototype scientist driven by the desire to know how things work. You're a smart girl which I happily made use of when I was stuck again. We go back quite a bit since the beginning of the Master's. I am happy that you also decided to continue your PhD here. I think back also at times outside of the lab where we enjoy ranting on about struggles or singing along to Of Monsters And Men. Thank you for supporting me until the very end, paranimf. Good luck to everything you are going to do in the future. **Sara**, we shared the small lab most of the time with just the two of us. Excellent moment to practice some German, yet often very disappointing from my side. Thank you for the memories. **Erik** and **Sebas**, the two of you cheer up the lab with your presence. All the best to the 2 of you. **Thomanai**, welcome to the group and good luck during your PhD. **Vida, Irati** and **Bárbara**, the nanobody-therapy trio. Good luck with the rest of your PhD's and hope you'll do well.

I also had the pleasure to share the building with the other research groups in Bloembergen. **Miranda**, you've grown a lot during the time that you were here. Truly developed as you like to term "strong independent woman" which is absolutely true. I'm happy that you managed to make the move to het Haagje as your heart lies there. All the best to you. **Rhythm**, you've started recently in the department sparkling like a little ball of lightning. I'm sure you've got a great future ahead of you and good luck with your PhD. **Maik,** i am sure that with the supervision of Markus you'll get the job done. Good luck with you pindakaas filled PhD. The other branch of NMR also contributed a lot during my time here. **Velten**, you're incredible dry sense of humor still haunts me today. It has been great the get to you know you. Good luck in the rest of your life. **Clara,** a true blauw vinger. Every time you would bring your mug, people would make comments about it, but we stayed loyal to our home town. All the best to you and your family. **Heyi**, you are absolutely unique. Incredibly laid back with the most intimidating walk. You never put yourself in the spotlight, yet we all got to know you bit by bit. **Ulric** (Le Paige), the amount of things you know something about is staggering. Doesn't matter the topic, you always know something to inform others with. Thank you for the waffles! **Vincenzo**, (man, I really hope I wrote your name correctly this time) an absolute powerhouse in the lab. Clara brought you up the speed but you're running the show now. Good luck with the rest of your PhD. I also had the pleasure to meet a great group of Post docs from the computational group. **Adrien, Li, Zeynep, Anna, Mickael, Jörg,** thank you all for a lovely time and I wish you all the best in life. **Gydo**, the main thing you taught me was how not to eat kwark which is a more important then you might think. **Cunliang** (Mr. Geng) thank you for advertising Chinese new year's every time with so much passion. You managed to convince many to join in for dumplings! **Panagiotis** (Panos), you have been a great support for a large group of people here. You don't mind being more to yourself but that doesn't mean you are not there for people. You're confident in what you do and know, stay true to that. **Francesco**, life is always a thumbs up with you. Prego mille for the fun and introducing me to the Dutch Lunch gallery. **Siri,** you've been a lively addition to the group. Your laughter is heard in the entire top floor which brightens the mood. I'll keep bringing kakis for lunch though. **Zuzana**, another cheerful addition to the computational group. Within a short time you established yourself by Slavic squads and "s"partying. **Jorge**, I wish you all the best in whatever comes your way and good luck in your future. **Charlotte**, I hope that you'll have a great time during your PhD and hope that you're happy along the way. **Brian, Rodrigo and Manon**, I hope that you will have a lot of joy during the time that you are here and all the best to you.

I also supervised several Bachelor and Master students. To **Michelle, Marco and**

Ruben, I would like to say thanks for the patience you had with me. You were the first students I supervised so I can imagine it was not completely adequate at all times. I was also just learning. I wish you all the best in your future studies. **Jeroen**, you were one of the fasted learning students I know. I literally have to show you a new experiments once and you could repeat it correctly without me having to look over your shoulder. You did great work and I was happy to supervise you. **Apfrida**, you really came at a great time. You pushed the projects further with great enthusiasm allowing them to progress at a higher pace. It must not always have been easy as your part of the project was in between 2 chapters of this thesis, making it quite challenging at times. I really enjoyed working together with you and at least I didn't scare you off from doing a PhD as you are pursuing your own at the moment. I'm sure you'll do well!

Finally, there are some additional people in the lab that one way or another influenced me. **Eugene**, since my Master's you given me instructions and advice on microscopy sample prep and imaging of these slides. I'm deeply impressed with what you are able to do and I would like to thank you for those sharing bits of knowledge with me. **Martin**, you're a scientist pur sang. The immense broad interest is a blessing to bounce some ideas off on. I hope that you'll be able to do the experiments you want to do and on your own quest in science. **Phebe**, thank you for everything that you do for me in the lab. It's probably much more then I actually realize. Moreover, being able to reenact Inside Out together with you was hilarious. You have a talent for acting!

In my Master's, I was lucky enough to meet an amazing bunch of people. The Westsiders were hot back then and are still a thing! Together with Cátia and Katerina, **Peter Jan** (PJ), **Rosanna**, **Robin** and **Maria**, we had a lot fun. Katerina, PJ and Robin decided to stick around a bit loner, meaning that we could still see each other on a daily basis. It's been great to do this journey together with you. Rosanna, I hope that you stay the happy person you are and wish you all the best. Maria, unfortunately we don't get to see you that often anymore, but when we do it's like old times. Good luck in Greece and wish you all the best!

Luckily, PhD life is not all about work work work. You need to take moments for yourself to enjoy other things in life that also allow yourself to let go. I'm happy I was persuaded to take part in an "open lesson" at the local dance school which got me hooked pretty fast. For years, the Afromagies were self-proclaimed dance group that performed live every Tuesday. **Doenja, Elisa** en **Vuza**, thank you for all the amazing moments we shared together. I really needed these moments to forget about the work for a moment to make my head clear and regain tons of energy. We also started to meet outside of the lessons for game nights where Jenga, off course, is the closing event of the evening. Where sweaty armpits and trembling hands are nothing new and Vuza just keeps on recording! Just thinking about it gives me stress. You are all great people and I wish you all the best in life.

Via Cátia I was also introduced to some truly amazing people that are also known on the 5th floor in Kruyt. **Dennis**, you managed to distract me enough at times from the work. Thank you for the lovely lunches in the parks and board games nights. Good luck for your future! To **Marta** and **Inês**, you are great friends to Cátia. I would like to say thank you for all the wonderful moment we shared together and all of the advice that you gave me. Your experience helped me. But more importantly were the moments to relax. I loved the moments we shared as a group, whether it is on Kingsday, a scorching hot summerday in the local pool on the balcony of Inês or just on anyones couch playing boardgames. I hope we will have many more great moments together in the near future. **Daniel** and **Roland**, the

two of you are amazing. We can have long and deep conversation about almost anything. I'm not great at small talk, but being able to discuss anything in depth is such a joy. Daniel good luck with your PhD, it will be tough at times but you're a smart guy and I'm sure you can do it. Roland, you are incredibly wise and manage to be supportive when things go south. A appreciate the calmness and kindness in both you. Thank you for our friendship. And finally, the small big man **Tiago**. It has been a pleasure to see you grow bit by bit. Even though at times my beard might scare you, I'm really not that scary behind it. You are very observant from a young age, maybe somewhat of a scientist yourself? Not so strange knowing who the parents are. Discover life as it comes.

I was also fortunate enough to love from a foreign place, which means that family visits would automatically be relaxing holidays. **João** and **Maria**, thank you for taking me into your home for already countless of visits. For me, spending away from here is a great moment to calm the nerves and get back to senses. Especially after our runs or hikes into the mountains. I consider myself lucky to have you as parents-in-law and hope that we will all one day visit Madeira all together. Thank you for all your support. **Tia Goretti** and **Tia Isabel**, thank you for your love and support. I hope that my sogra did not scare you too much after my first trip to the beach! Such sandy beaches. **Carla, Andreia** and **André** I would also like to say thanks for being part of this journey with me and for all the dinners you prepared. I'm still full from last time!

Pap en **mam**, bedankt voor alle steun die jullie mij deze tijd gegeven hebben. Ik kan me voorstellen dat het voor jullie niet altijd even makkelijk is om je voor te stellen waar ik nu precies door heen ga. Gelukkig waren jullie heel geduldig met me. Bedankt voor alles wat jullie voor me doen. **Marleen** en **Ellen**, mijn twee grote zussen. Ik wil ook jullie bedanken voor de steun die jullie me gegeven hebben. Jullie zijn beide een voorbeeld voor me geweest, zowel tijdens het opgroeien als nu. Ik mag me gelukkig prijzen met zussen als jullie. **Ron** en **Bart,** ik moet jullie bedanken met name voor de afleiding die jullie mij geschonken hebben. Soms moet je gewoon even iets anders doen, balletje tikken, overgooien, rond-de-tafel, you name it. Bedankt voor alles. **Linde**, mijn slimme meid. Ik ben heel erg trots om je oom te zien en heb met veel plezier je zien opgroeien. Ik hoop dat je net zo graag de wereld wilt ontdekken als ik! **Oma An** en **Oma Mannie**, mijn omaatjes. Jullie zijn beide sterke vrouwen, zeker vrouwen om trots op te zijn.

And now the time has come to thank the person that contributed the most during this journey. Sometimes it seems easy to say what is on your mind, but ends up being the hardest to express in words. **Cátia**, I cannot express how thankful I am for everything that you did for me. It all started when we found each other here in Utrecht. Moving from Groningen to Utrecht for my Master's was the smartest thing that I could have done. You've become the foundation of my life and gave me the support I needed during this period. I cannot thank you enough for all your help and patience as I'm sure I wasn't always the easiest to handle. I relied a lot on you. You are the kindest, sweetest and most caring person I know and you managed to motivate me when things were tough. We often say "And if you are in love, you are the lucky one". These words have proven again and again to be more than true. I'm thankful for all the love and affection that you have always given me. I cannot wait to discover what life has to offer with you. This chapter finishes here, but our journey will continue. Amo-te imenso paixão.

This is the end, hora est!
Reinier

