

N-Glycosylation as determinant of epidermal growth factor receptor conformation in membranes

Karol Kaszuba^{a,1}, Michał Grzybek^{b,c,1}, Adam Orłowski^a, Reinis Danne^a, Tomasz Róg^a, Kai Simons^{d,2}, Ünal Coskun^{b,c,2}, and Ilpo Vattulainena,e,2

^aDepartment of Physics, Tampere University of Technology, FI-33101 Tampere, Finland; ^bPaul Langerhans Institute Dresden of the Helmholtz Centre Munich at the University Clinic Carl Gustav Carus, TU Dresden, 01307 Dresden, Germany; German Center for Diabetes Research (DZD e.V.), 85764 Neuherberg, Germany; ^dMax Planck Institute for Molecular Cell Biology and Genetics, 01307 Dresden, Germany; and ^eCenter for Biomembrane Physics (MEMPHYS), University of Southern Denmark, 5230 Odense, Denmark

Contributed by Kai Simons, February 24, 2015 (sent for review April 10, 2014; reviewed by Helmut Grubmüller and Paul M. P. Van Bergen en Henegouwen)

The epidermal growth factor receptor (EGFR) regulates several critical cellular processes and is an important target for cancer therapy. In lieu of a crystallographic structure of the complete receptor, atomistic molecular dynamics (MD) simulations have recently shown that they can excel in studies of the full-length receptor. Here we present atomistic MD simulations of the monomeric N-glycosylated human EGFR in biomimetic lipid bilayers that are, in parallel, also used for the reconstitution of full-length receptors. This combination enabled us to experimentally validate our simulations, using ligand binding assays and antibodies to monitor the conformational properties of the receptor reconstituted into membranes. We find that N-glycosylation is a critical determinant of EGFR conformation, and specifically the orientation of the EGFR ectodomain relative to the membrane. In the absence of a structure for full-length, posttranslationally modified membrane receptors, our approach offers new means to structurally define and experimentally validate functional properties of cell surface receptors in biomimetic membrane environments.

EGFR | lipids | MD simulation | lipid-protein interaction | proteoliposomes

Receptor tyrosine kinases (RTKs) are cell surface receptors that receive and transduce signals mediating a variety of critical cellular processes, including cell growth, migration, proliferation, differentiation, and apoptosis. Among the many RTKs, the most studied is the epidermal growth factor receptor (EGFR), not least because of its involvement in the development and progression of epidermoid cancers and its resulting importance as a target for antineoplastic therapies.

Structurally, the EGFR consists of the ectodomain (ECD) (further subdivided into four subdomains, DI-IV), the transmembrane domain (TMD), and the intracellular tyrosine kinase domain (TKD). Ligand binding induces conformational transitions of the ECD that stabilize receptor dimerization, culminating in the activation of the intracellular TKD and subsequent propagation of the activation signal (1). To prevent receptor activation and signaling in the absence of ligand, the structurally tethered ECD of monomeric EGFR blocks the intrinsic capacity of the TMD and the intracellular TKD to dimerize (2). Ligand binding is believed to release the self-inhibitory tether and facilitate receptor oligomerization and activation (3-6). A detailed understanding of the structural regulation of the intact full-length receptors in their native membranes promises to reveal the molecular basis for receptor regulation (7); however, the methodological limitations associated with crystallizing transmembrane proteins, together with the high flexibility of the full-length receptor, have prevented high-resolution crystallographic analysis.

To fill this gap, extensive atom-scale molecular dynamics (MD) simulations were recently performed to elucidate the structural dynamics of the EGFR in a two-component lipid bilayer (8). These studies suggest a large interfacial contact area between the membrane and the ecto- and intracellular domains of the unliganded monomeric and dimeric receptors. However, experimental data at the single-molecule level stipulate a greater distance between the ECD of the receptor and the membrane in the absence of ligand (9, 10).

To resolve this discrepancy, the present study considers two key parameters that were not included in previous simulations of the monomeric EGFR: N-glycosylation of the ectodomain of EGFR that contributes up to 50 kDa of the total molecular weight of ~178 kDa (11, 12), and a lipid environment that prevents the EGFR from ligand-independent, and thus aberrant, activation (13). Comparing glycosylated and nonglycosylated EGFR, we find that N-glycosylation is critical for the conformational arrangement of the ECD subdomains DI-IV and their interfacial contact area with the membrane. In our simulations, only the N-glycosylated receptor adopts a conformation that is in good agreement with previous experimental FRET studies for the distance between the EGFR ECD and the membrane (9, 10). Experimental ligand and antibody binding assays on EGFR proteoliposomes provide validation of the MD simulations.

Results

To follow the dynamics of membrane-embedded monomeric EGFR, we have performed extensive atomistic MD simulations of the near-full-length receptor [lacking the unstructured C-terminal

Significance

Structural analysis of growth factor receptors in their membrane environment is key for understanding their functions that are vital to the development and survival of organisms. High structural flexibility and posttranslational modifications of the fulllength receptors, however, hinder structural analysis at high resolution. Here, we used atomistic molecular dynamics simulations and biochemical experiments with proteoliposomes to elucidate the role of N-glycosylation with regard to the structural properties of the human epidermal growth factor receptor (EGFR). We find that N-glycosylation critically determines membrane interactions and structural arrangement of the ligand-binding EGFR ectodomain. This combined approach provides new means to structurally explore and experimentally validate functional properties of cell surface receptors and test therapeutic agents, such as monoclonal antibodies.

Author contributions: K.S., Ü.C., and I.V. designed research; K.K. and M.G. performed research; K.K., M.G., A.O., R.D., and T.R. contributed new reagents/analytic tools; K.K., M.G., Ü.C., and I.V. analyzed data; and K.K., M.G., K.S., Ü.C., and I.V. wrote the paper.

Reviewers: H.G., Max Planck Institute for Biophysical Chemistry: and P.M.P.V.B.e.H., Utrecht University.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

¹K.K. and M.G. contributed equally to this work.

²To whom correspondence may be addressed. Email: simons@mpi-cbg.de, coskun@plid.de, or ilpo.vattulainen@tut.fi.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10. 1073/pnas.1503262112/-/DCSupplemental

tail of the kinase domain, as in previous MD simulations (8)]. N-glycosylation is required for EGFR trafficking, efficient ligand binding, and receptor activation (12, 14-17), suggesting this modification confers an additional informational level to the underlying polypeptide structure (18), required for achieving a functional receptor state. Although number, sequence, size of branches, and fucosylation of glycans vary between specific proteins and cell types, Man₃GlcNAc₂ residues constitute the cell type-independent minimal core for all N-glycans attached to proteins, critical for protein folding in the endoplasmic reticulum (19, 20). Here, we compared the receptor conformation either lacking or containing Man₃GlcNAc₂ glycan residues at 10 welldescribed positions of the ectodomain (21, 22) (SI Appendix, Fig. S1).

The receptor was embedded in a membrane composed of a ternary lipid mixture of 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), N-stearoyl-D-erythro-sphingosylphosphorylcholine (SM), and cholesterol, a plasma membrane mimetic composition that recapitulates as closely as possible the experimental conditions for biochemical reconstitution of the EGFR in proteoliposomes (SI Appendix, Table S1) (13). In synthetic membranes, similar ternary lipid compositions give rise to two immiscible fluid membrane phases: a DOPC-enriched liquid-disordered and a cholesterol/SMenriched liquid-ordered phase (23). Using this lipid composition for EGFR proteoliposomes, we previously showed that the presence of both phases is critical to prevent ligand-independent kinase activation, whereas membranes with low cholesterol levels allowed ligand-independent receptor activation (13), in agreement with aberrant EGFR activation in cellular membranes on cholesterol depletion (24).

Atomistic simulations of membrane receptors are challenging because the time scales of conformational changes are relatively slow, and therefore adequate sampling of receptor conformation is an important concern. In this work, we have accounted for this issue by carrying out a number of complementary simulations. For both the glycosylated and the nonglycosylated receptors, the MD simulations covered 1,000 ns (together these constitute "simulation 1"; for PDB files, please see Datasets S1 and S2). To improve sampling, we next started from different (independent) initial conditions and repeated 1,000-ns runs for both the glycosylated and the nonglycosylated EGFR ("simulation 2"; for PDB files, please see Datasets S3 and S4). In addition, the intermediate configurations of the first simulations at 300 and 1,000 ns were used to consider how the glycosylated EGFR evolves toward the nonglycosylated structure after the removal of glycans.

Conformational Arrangement of the EGFR Ectodomain. In all configurations (nonglycosylated and glycosylated), the ECD of the EGFR underwent large-scale conformational changes within the first 500 ns (Fig. 1) as a consequence of the hinge-bending domain motions at two different regions of the ECD (SI Appendix, Fig. S2).

The first of the identified hinges (H1) was formed by a short loop (amino acid residues 615-621) connecting the ECD to the TMD of the receptor. The second hinge (H2) encompassed the residues 230-239 and 259-261 of subdomain DII. Within the first nanoseconds, the ECD bending motion around hinge H1 resulted in the formation of a large contact interface area between subdomain DIV and the membrane, in all configurations (Fig. 1 C and D). In the nonglycosylated receptor, hinge H2 caused a rotation of subdomain DI toward the membrane. Once stably attached, subdomain DI buried ~1,000-2,000 Å² of the solvent-accessible surface area (SASA) in the interface with lipids (Fig. 1 B and D).

The presence of Man₃GlcNAc₂ glycan residues significantly altered the relative arrangement of the individual receptor subdomains and their alignment on the membrane (Fig. 1, SI Appendix, Fig. S3, and Movies S1 and S2). The glycan moieties attached at positions

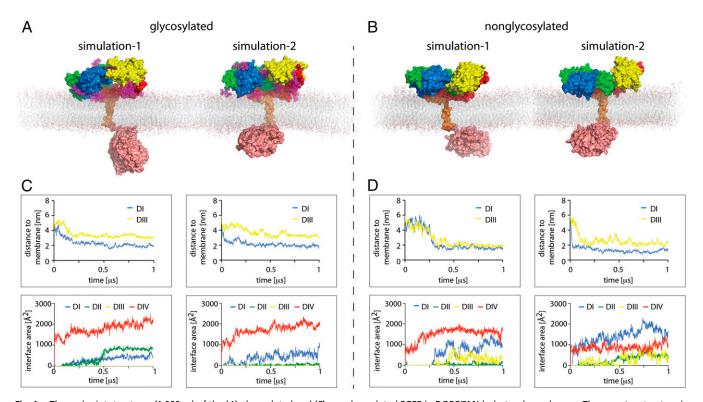


Fig. 1. The endpoint structures (1,000 ns) of the (A) glycosylated and (B) nonglycosylated EGFR in DOPC/SM/cholesterol membranes. The receptor structure is color-coded throughout the panels as follows: subdomains DI (blue), DII (green), DIII (yellow), and DIV (red); TM domain (orange); intracellular TKD (salmon); glycans (purple). (C and D) The distances of the center of mass of subdomains DI and DIII to the membrane (Upper) and membrane interface area of ECD subdomains DI-DIV (Bottom), for (C) glycosylated and (D) nonglycosylated EGFR. For corresponding PDB files, see Datasets S1-S4.

N151 (DI), N172 (DII), N389 together with N420 (DIII), and N504 (DIV) were found to behave as molecular cushions, lifting these domains away from the plane of the membrane and changing their relative arrangement to each other. The elevation of the EGFR structure above the membrane surface relates particularly to subdomains DI and DIII, which together contribute to ligand binding (4, 25). The glycan moiety at position N151 in subdomain DI reduced its rotation around H2, elevating subdomain DI and drastically decreasing its interface area with the membrane (Fig. 1 C and D and SI Appendix, Table S2). Notably, the Man₃GlcNAc₂ residues in DIII entirely prevented membrane contacts of this domain (Fig. 1C). Moreover, the glycan residue at position N328 functioned as a spacer, defining the relative arrangement of subdomains DI and DIII to each other (Fig. 1A and SI Appendix, Fig. S3). Interestingly, N328 is one of the four N-glycosylation sites within subdomain DIII that are probably sufficient to induce a ligand-binding competent receptor conformation (12).

To further test the contribution of glycans on ECD arrangement, we performed a control simulation starting at the 1,000-ns configuration of glycosylated simulation 1. After removal of all glycans, the receptor would be expected to convert toward the nonglycosylated end-structure. The receptor evolved gradually, but the changes observed during additional microsecond MD simulations were limited, presumably as a consequence of already stabilized membrane-protein interactions. We therefore repeated this approach but removed all glycans after 300 ns, a point at which membrane-contact sites had not yet been stabilized. Interestingly, within the next 900 ns, the in silico deglycosylated receptor evolved toward the end configuration of the nonglycosylated receptor simulations, confirming the importance of N-glycosylation in EGFR ECD arrangement. In the absence of glycan residues, DI and DII now formed again a contact site, and DIII steadily approached the membrane until it formed an interface with the lipid bilayer (SI Appendix, Fig. S4).

Experimental Validation of the MD Simulations for the ECD Conformation.

In the crystal structures of the enzymatically (partially) deglycosylated and ligand-stabilized ectodomain dimer, EGF is bound between subdomains DI and DIII (4, 5). However, according to our simulations, the ligand-binding site at DI in the non-glycosylated monomeric receptor would be masked by the membrane, constraining placement of a ligand at this position without significant alteration of the ECD conformation (Fig. 2D). This is not the case for the glycosylated receptor, where the ECD adopts an elevated and more defined arrangement. These observations motivated us to experimentally test whether ECD glycosylation has an influence on ligand-binding properties in EGFR proteoliposomes. In all our ligand-binding assays, equilibrium dissociation constants of the ligand were in the range of ~2 nM (Fig. 2B, Inset), representing the low-affinity receptor class (26).

To experimentally validate the structural rearrangement of the EGFR subdomains predicted by our simulations, we compared the maximum number of binding sites (B_{max}) of two monoclonal antibodies at saturating conditions: C225 (cetuximab) and 2E9 (27, 28). The therapeutic monoclonal antibody C225 recognizes an epitope on DIII that completely inhibits ligand binding. In our in vitro proteoliposomal assay, C225 binding to the receptor was independent of receptor glycosylation (Fig. 2B), and we could recapitulate the complete inhibitory action of C225 on EGFR ligand binding (Fig. 2C). This is in very good agreement with our MD simulations that suggest no dependence of C225 epitope accessibility on EGFR glycosylation (Fig. 2D). We repeated the binding assay with the monoclonal antibody 2E9 that specifically inhibits EGF binding to the low-affinity class of receptors (~95-98% of all receptors present at the plasma membrane). Unlike the original paper (29), we limited the incubation time with the 2E9 antibody to 1 h, precluding complete

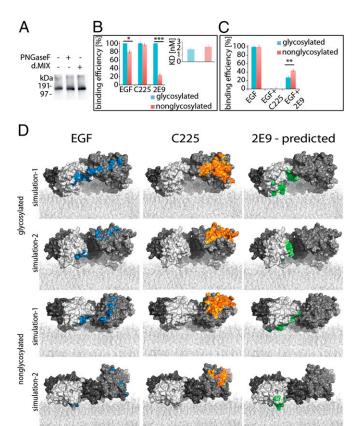


Fig. 2. (A) Western blot analysis of enzymatically (partially) deglycosylated EGFR (postreconstitution in proteoliposomes), using either peptide-Nglycosidase F or deglycosylation mix. (B) Relative binding efficiencies at B_{max} of EGF (150 ng/mL), mAb-C225 (1 µg/mL), or mAb-2E9 (1 µg/mL) to glycosylated or deglycosylated EGFR proteoliposomes. Insert shows EGF dissociation constant at equilibrium (KD) before and after deglycosylation of reconstituted EGFR. (C) Relative EGF binding efficiency at $B_{\rm max}$ after preincubation of the proteoliposomes with mAb-C225 (1 µg/mL) or mAb-2E9 (1 μg/mL). EGF binding in the absence of antibodies was used as control and normalized to 100%. All measurements were performed in three independent experiments, each of them in either duplicates (K_D) or triplicates (antibody and ligand-binding efficiencies at $B_{\rm max}$). (D) Representation of EGFR ectodomain after 1,000 ns, with highlighted EGF binding sites colorcoded in blue (based on crystal structure 1NOL) (4), mAb-C225 in orange (based on 1YY8 and 1YY9) (27), and the predicted mAb-2E9 epitope in green (based on the data presented in this work).

receptor inhibition. Enzymatic deglycosylation of the receptor reduced the binding efficiency of 2E9 by \sim 75% (Fig. 2B), also translating into less-efficient inhibition of ligand binding (Fig. 2C). Previous work suggested that 2E9 would recognize an epitope located on DI (30), but no structural evidence is available. According to our experimental data, the 2E9 epitope is partially obscured in the enzymatically deglycosylated EGFR structure (Fig. 2B).

To identify residues that are likely part of the 2E9 epitope, we used the atomistic simulation data and performed SASA calculations for the ECD and averaged the results over the last 50 ns. The average $\Delta SASA$ ($\Delta SASA = SASA_{glyco} - SASA_{nonglyco}$) was used to map those residues with major changes in solvent accessibility within all four major simulations. Interestingly, the residues we identified in DI (with a cutoff filter of $\Delta SASA > 0.4 \text{ nm}^2$) are part of, or located in the vicinity of the reported ligand-binding site in DI (4) (Fig. 2D and SI Appendix, Fig. S5). An alternative scenario is the direct involvement of glycosylation moieties in the 2E9 epitope. This, however, is unlikely, as branching and length of

Kaszuba et al. PNAS Early Edition | **3 of 6**

glycosylation patterns are cell type-specific, whereas 2E9 inhibits low-affinity EGFRs in various cell types (29, 31, 32), as well as the recombinant insect SF+ cell protein product in our experiments. Overall, our data support the view that DIII is the primary ligand-binding site, and the 2E9 antibody inhibits low-affinity EGF binding through preferential interactions with DI.

Conformational Coupling Across the Membrane. The EGFR transmembrane domain has been previously described to be critically involved in the communication between the ECD and the intracellular TKD. Therefore, the possible regulatory function of the TMD has received substantial attention (2, 8, 33–38). Two GxxxG dimerization motifs present on both ends of the TMD are supposed to regulate the association of ligand-free and ligand-bound dimers (2, 8). In our MD simulations, we did not observe glycan-based structural changes in the C-terminal dimerization motif (⁶³⁷ALGIG⁶⁴¹) (*SI Appendix*, Fig. S6 *A* and *B*), whereas the N-terminal motif (⁶²⁵GMVGA⁶²⁹) was flexible and eventually even unfolded in nonglycosylated EGFR simulation 1.

N-terminal TMD dimerization is also required as a regulatory step for asymmetric dimer formation, and hence activation of the intracellular TKD, a process guided through membrane proximal, intracellular juxtamembrane segments A and B (JM-A and JM-B). Upon N-terminal TMD dimerization, the JM-A segments interact with the membrane and form an α -helical, antiparallel homo-dimer (8, 39, 40). The α -helical structure of the JM-A dimer has been confirmed in NMR and MD simulation studies of TMD-JM-A fragments (2, 8, 41) and in MD simulation of the active EGFR dimer (8).

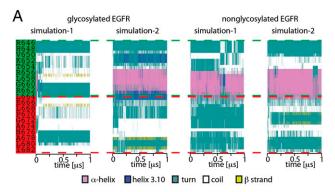
In both of our simulations of the nonglycosylated receptor, the JM-A segment is primarily α-helical (Fig. 3). Interestingly, the helical property of the JM-A segment strikingly correlates with the formation of a rapidly evolving and stable contact interface area of the TKD with the membrane (Fig. 3B, SI Appendix, Fig. S7, and Movie S2). In the glycosylated EGFR, this feature can be observed in simulation 2 only. In simulation 1, the TKD does not form a large membrane contact interface with the same kinetics, although its membrane interactions are stable and increase gradually (SI Appendix, Fig. S7, and Movie S1). As a consequence, the initially α-helical JM-A segment unfolds promptly at the beginning and remains unfolded throughout the entire simulation. We also observe that the membrane interaction of the TKD is not restricted to a preferred orientation.

Discussion

High-performance atomistic MD simulations of the full-length EGFR embedded in membranes in combination with experimental reconstitution systems are a promising approach to bridge the gap between biochemistry and classical structural methods. With this multipronged approach, we have evaluated the effect of N-glycosylation of the EGFR in a defined lipid environment that promotes ligand-dependent EGFR activation.

Unique among posttranslational modifications, glycosylation is immensely diverse, with different cell lines exhibiting distinct patterns of protein glycosylation, largely as a result of the expression of varying repertoires of glycosidases and glycosyltransferases (20). Hence, the sugar sequence, number and size of branches, and fucosylation of additional glycans vary between specific proteins and cell types.

The heterogeneity, considerable size, and flexibility of glycan residues hinder protein crystallization. For the human EGFR, among other proteins, this problem can be bypassed by recombinant production of the receptor in insect cells (27, 42, 43) in which glycosylation sites, ligand binding, and receptor activation properties are not compromised, although glycan branching and branch length are significantly reduced. Alternatively, the receptor is expressed in mammalian cells and then partially enzymatically



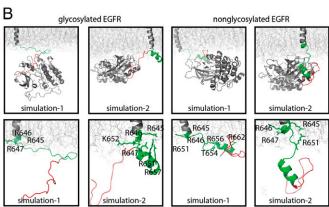


Fig. 3. (A) Secondary structure changes of the intracellular JM-A and JM-B segments, color-coded in green and red, respectively. (B) The structure of the EGFR intracellular domain and hydrogen bonding between JM-A and lipids at the 1,000 ns endpoints of the respective simulations. Amino acid residues that interact with lipids for more than 50% of the simulation time are highlighted as stick representation. Corresponding residue occupancy times can be found in SI Appendix, Tables S3 and S4.

deglycosylated before crystallization (4, 5). The resulting polypeptide retains some *N*-glycosylation residues, possibly because they are structurally inaccessible to enzymatic cleavage in the folded receptor complex. However, even when present, glycan structures are only partially resolved in crystal structures of the EGFR ectodomain (4, 5, 27, 42, 43). Therefore, the structural consequences of glycosylation, and the resulting functional effects on receptor regulation are difficult to assess.

On the basis of our MD simulations and experiments, glycosylation critically determines the structural arrangement of the ECD, and in particular, of the ligand-binding domains DI and DIII. Our data confirm the notion that DIII is the major site of ligand binding, whereas DI has only a minor contribution (30, 44). Additionally, DIII is known to bind the ligand when expressed as soluble protein (25) or upon proteolytic cleavage of the ECD (45); antibodies and nanobodies that bind to DIII [C225 (cetuximab), IMC11F8, VHH 7D12] compete with EGF for the binding site on this subdomain (27, 42, 46).

To further interrogate the credibility of the simulated structures, we compared the computed protein membrane distances to the previously reported simulations of the monomeric EGFR (8), as well as to experimental data (9, 10). Although our simulations of the nonglycosylated receptor are in good agreement with the previous simulations (8), significant differences in the arrangement of the ECD subdomains DI and DIII are observed for the glycosylated receptor (*SI Appendix*, Table S2). Here, the presence of glycans propelled DI, and especially DIII, away from the membrane, up to ~1.9 and 3.2 nm, respectively (the distance computed along the membrane normal from the center of mass

of the ECD subdomain to the average position of phosphorous atoms in lipid headgroups) (Fig. 1C and SI Appendix, Table S2). FRET studies with N-terminally fluorescence-tagged EGFR suggested that the distance between the fluorescent label and the membrane ranges from ~ 6.2 to 8 nm (9, 10). This is in discrepancy with the simulations of the nonglycosylated receptor, where the N terminus of the receptor is located ~3.1-3.5 nm away from the membrane (SI Appendix, Table S2). In comparison, the N-terminal distance from the membrane for the glycosylated receptor is 4.5 nm. Assuming that a fluorescent dye or probe would contribute additionally at least 1 nm, the total distance would increase to 5.5 nm, which is in much better agreement with the published experimental data of 6.2-6.4 nm and 8 nm for the acyl carrier protein (9) and YFP-tagged EGFR (10), respectively. The difference between simulations and the experimental single-molecule data is still discernible, but is expected since we have considered the Man₃GlcNAc₂ core glycosylation only, accounting for about 30% of total glycans of the EGFR expressed in mammalian cells. Moreover, the plasma membrane also contains glycolipids known to interact with the EGFR ECD (13, 47, 48), which may also increase the observed distances at the cellular level.

Critically, however, for our understanding of receptor activation, the coupling mechanism between the ECD and the intracellular kinase domain across the biological membrane remains ambiguous. Very recently, Arkhipov et al. presented MD simulations of the ligand-stabilized and glycosylated human EGFR dimer that lacks the intracellular juxtamembrane and kinase domains (49). Compared with their previous simulations of the liganded, but nonglycosylated full-length EGFR dimer (8), the ECD now significantly interacts with the membrane. Truncation of the ectodomains, in contrast, leads to aberrant dimerization and activation of the EGFR kinase domain (2). The ill-defined linkage mechanism between ligand binding and stabilization of the active kinase domains stems, at least partially, from methodological limitations resulting from solubilized receptors being studied in detergent micelles, augmenting flexibility of the membrane proximal sequences; that is, the kinase domains of ligandbound dimeric receptors can adopt flexible conformations (50) that correlate with the active or inactive state of EGFR dimers (51). It is therefore tempting to conclude that the membrane itself holds the key for understanding domain coupling (7).

Eukaryotic cells tune the composition of their membranes through directed lipid sorting along the secretory pathway and selective lipid transport across the bilayer. As a consequence, the physicochemical properties of their membranes vary significantly throughout the cell (52), exemplified by the plasma membrane being highly enriched in cholesterol (~35–40 mol%) and sphingolipids, both critically regulating membrane fluidity and thickness (53). Upon cholesterol depletion, the EGFR kinase domain is activated in a ligand-independent manner (24, 54). A very similar effect was observed in synthetic reconstitution experiments: ligand-dependent EGFR activation was observed only in bilayers with high cholesterol and sphingomyelin levels, while reconstitution of EGFR into bilayers enriched in low melting temperature phosphatidylcholine led to aberrant receptor activation (13). Previous MD simulations and NMR studies have not accounted for the compositional and physicochemical particularities of the plasma membrane. Occasionally used synthetic lipids such as 1,2-dipalmitoyl-sn-glycero-3-phosphocholine and 1,2-distearoyl-sn-glycero-3-phosphocholine do have high transition temperatures (41 °C and 55 °C), but these do not mimic the fluidity of natural membranes in the absence of cholesterol. In contrast, short-chain lipids such as 1,2-dimyristoyl-sn-glycero-3phosphocholine form thin and homogenous bilayers, and therefore are expected to energetically constrain the TMD, as described for synthetic transmembrane helices in membranes of varying thickness (55).

For our MD simulations, we have used a lipid composition that is as close as possible to the ternary lipid mixture that was used in the reconstitution experiments, which are themselves a reasonable (although dramatically simplified) mimic of the three key lipid types (saturated, unsaturated, cholesterol) present in live cell plasma membranes. The advantage of this approach is that it provides mutual validation between computational and biochemical experiments. Although the MD simulations of the glycosylated receptor lead to a highly reproducible structural arrangement of the EGFR ECD on the membrane, the properties of the juxtamembrane fragments and orientation of the TKD are flexible throughout all simulated systems. In the MD simulations of Arkhipov et al. (8), the TKD and JM-A fragment are attracted to the membrane by the presence of negatively charged phosphatidylserine. In our systems, the TKD and JM-A fragment still are capable of interacting with the membrane (Fig. 3 and *SI Appendix*, Fig. S7, Tables S3 and S4), although it contains neutral lipids only. In contrast to the well defined ECD arrangement, the intracellular modules interact with the membrane without any preferred orientation. The presence of negatively charged lipids could possibly change this into a guided process. Moreover, the stability of the α-helical JM-A segment seems to depend on a critical threshold of the membrane contact area interface of the entire TKD, rather than the direct interactions of this fragment with the lipids. Unlike the other simulations, in simulation 1 of the glycosylated receptor, JM-A unfolds promptly and remains unfolded throughout the simulation, which correlates with the limited formation of the membrane interface. This finding, together with the previously described electrostatic membrane interactions with lipids, highlights the capacity of the intracellular JM-A domain to sense and interact with various lipid environments, possibly in an electrostatic switch-like mechanism (6, 8, 56, 57). It should also be noted that the orientation of the TKD and the helicity of the JM-A segment do not seem to have an influence on the properties of the kinase active site of the TKD. Independent of glycosylation state or membrane association, the TKD retains the previously described and characterized features of the nonactive conformation, with the activation loop inserted in the kinase active site and the inactive orientation of the helix αC (39, 51) (SI Appendix, Fig. S8).

In summary, glycosylation of the EGFR appears to be critical for the conformational arrangement of the ECD and its membrane contact sites, whereas the influence on TMD properties, JM-A segment, and the orientation and membrane association of the TKD are not significantly affected. Our approach to study the full-length, posttranslationally modified EGFR by high-performance MD simulations supplemented with proteoliposomal reconstitution systems has the potential to become a comprehensive toolbox for studying membrane proteins in biomimetic lipid environments.

Materials and Methods

Reconstruction of EGFR Receptor Structure for Atomistic MD Simulations. A nearly full-length receptor (2–994) that lacks only the very C-terminal autophosphorylation tail (residues 995–1,186) was reconstructed from separately crystallized PDB structures: 1YY9 (27) for the extracellular domain and 2RGP/3GOP (40, 51) structures for the intracellular TK domain. The structure of the TM helix was built de novo. The fully reconstructed EGFR chain was subsequently embedded in a ternary lipid mixture (DOPC/SM/cholesterol). The lipid compositions of these systems were designed to match the previously reported in vitro reconstitution conditions (13). The receptor was *N*-glycosylated in silico with core Man₃GlcNAc₂ residues at documented sites (4, 5, 21, 22, 27, 42, 43).

We performed atomistic MD simulations for the glycosylated and the nonglycosylated receptor. Both systems were simulated for 1,000 ns. These studies are referred to in this work as simulation 1. In addition, we improved sampling by carrying out simulation 2, in which we repeated the 1,000-ns simulations for the glycosylated and nonglycosylated receptors by starting from different (independent) initial conditions. The results of these replicas were largely consistent with those of simulation 1, as discussed earlier (*SI Appendix*). A more detailed description of system construction, simulation conditions, additional simulation data as well as experimental methods are provided in *SI Appendix*.

Kaszuba et al. PNAS Early Edition | **5 of 6**

ACKNOWLEDGMENTS. We thank Cornelia Schroeder, Ilya Levental, Daniel Lingwood, Aleksander Czogalla, Donna Arndt-Jovin, and Thomas Jovin for critical reading of the manuscript. We acknowledge that the results of this research have been achieved using the Partnership for Advanced Computing in Europe (PRACE) Research Infrastructure resource Hermit based in Germany at the High Performance Computing Center Stuttgart (HLRS). CSC-Finnish IT Centre for Scientific Computing (Espoo, Finland) is acknowledged for additional computer resources. Support was provided by the Academy of Finland [project funding, Center of Excellence scheme (K.K., A.O., T.R., I.V.)], the European Research Council [Advanced Grant CROWDED-PRO-LIPIDS (T.R., I.V.), the Sigrid Juselius Foundation (K.K., A.O., T.R., I.V.), the Deutsche Forschungsgemeinschaft "Transregio 83" Grant TRR83 TP18, and the German Federal Ministry of Education and Research grant to the German Center for Diabetes Research (DZD e.V.) (M.G. and Ü.C.).

- 1. Ferguson KM (2008) Structure-based view of epidermal growth factor receptor regulation. Annu Rev Biophys 37:353-373.
- 2. Endres NF, et al. (2013) Conformational coupling across the plasma membrane in activation of the EGF receptor. Cell 152(3):543-556.
- 3. Burgess AW, et al. (2003) An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Mol Cell 12(3):541-552.
- 4. Ogiso H, et al. (2002) Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell 110(6):775-787.
- 5. Garrett TP, et al. (2002) Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. Cell 110(6):763-773.
- 6. Wang Y, et al. (2014) Regulation of EGFR nanocluster formation by ionic protein-lipid interaction. Cell Res 24(8):959-976
- 7. Bessman NJ, Freed DM, Lemmon MA (2014) Putting together structures of epidermal growth factor receptors. Curr Opin Struct Biol 29:95-101.
- 8. Arkhipov A. et al. (2013) Architecture and membrane interactions of the EGF receptor. Cell 152(3):557-569.
- Ziomkiewicz I, et al. (2013) Dynamic conformational transitions of the EGF receptor in living mammalian cells determined by FRET and fluorescence lifetime imaging microscopy. Cytometry A 83(9):794-805.
- 10. Kozer N, et al. (2011) Evidence for extended YFP-EGFR dimers in the absence of ligand on the surface of living cells. Phys Biol 8(6):066002.
- 11. Zhen Y, Caprioli RM, Staros JV (2003) Characterization of glycosylation sites of the epidermal growth factor receptor. Biochemistry 42(18):5478-5492.
- 12. Fernandes H, Cohen S, Bishayee S (2001) Glycosylation-induced conformational modification positively regulates receptor-receptor association: A study with an aberrant epidermal growth factor receptor (EGFRvIII/DeltaEGFR) expressed in cancer cells. J Biol Chem 276(7):5375-5383.
- 13. Coskun Ü, Grzybek M, Drechsel D, Simons K (2011) Regulation of human EGF receptor by lipids. Proc Natl Acad Sci USA 108(22):9044-9048.
- Wang XQ, Sun P, O'Gorman M, Tai T, Paller AS (2001) Epidermal growth factor receptor glycosylation is required for ganglioside GM3 binding and GM3-mediated suppression [correction of suppression] of activation, Glycobiology 11(7):515-522.
- 15. Soderquist AM, Carpenter G (1984) Glycosylation of the epidermal growth factor receptor in A-431 cells. The contribution of carbohydrate to receptor function. J Biol Chem 259(20):12586-12594.
- Yamabhai M, Anderson RG (2002) Second cysteine-rich region of epidermal growth factor receptor contains targeting information for caveolae/rafts. J Biol Chem 277(28):24843-24846
- 17. Tsuda T, Ikeda Y, Taniguchi N (2000) The Asn-420-linked sugar chain in human epidermal growth factor receptor suppresses ligand-independent spontaneous oligomerization. Possible role of a specific sugar chain in controllable receptor activation. J Biol Chem 275(29):21988-21994
- 18. Moremen KW, Tiemeyer M, Nairn AV (2012) Vertebrate protein glycosylation: Diversity, synthesis and function. Nat Rev Mol Cell Biol 13(7):448-462.
- 19. Kornfeld R, Kornfeld S (1985) Assembly of asparagine-linked oligosaccharides. Annu Rev Biochem 54:631-664
- 20. Helenius A, Aebi M (2001) Intracellular functions of N-linked glycans. Science 291(5512): 2364-2369
- 21. Sato C. et al. (2000) Characterization of the N-oligosaccharides attached to the atypical Asn-X-Cys sequence of recombinant human epidermal growth factor receptor. J Biochem 127(1):65-72.
- 22. Stroop CJ, et al. (2000) Characterization of the carbohydrate chains of the secreted form of the human epidermal growth factor receptor. Glycobiology~10(9):901-917.
- 23. Chiantia S, Ries J, Kahya N, Schwille P (2006) Combined AFM and two-focus SFCS study of raft-exhibiting model membranes. ChemPhysChem 7(11):2409-2418.
- 24. Ringerike T, Blystad FD, Levy FO, Madshus IH, Stang E (2002) Cholesterol is important in control of EGF receptor kinase activity but EGF receptors are not concentrated in caveolae. J Cell Sci 115(Pt 6):1331-1340.
- 25. Lemmon MA, et al. (1997) Two EGF molecules contribute additively to stabilization of the EGFR dimer. EMBO J 16(2):281-294.
- Ozcan F, Klein P, Lemmon MA, Lax I, Schlessinger J (2006) On the nature of low- and high-affinity EGF receptors on living cells. Proc Natl Acad Sci USA 103(15):5735-5740.
- Li S, et al. (2005) Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer Cell 7(4):301-311.
- 28. Goldstein NI, Prewett M, Zuklys K, Rockwell P, Mendelsohn J (1995) Biological efficacy of a chimeric antibody to the epidermal growth factor receptor in a human tumor xenograft model. Clin Cancer Res 1(11):1311-1318.

- 29. Defize LH, et al. (1989) Signal transduction by epidermal growth factor occurs through the subclass of high affinity receptors. J Cell Biol 109(5):2495-2507.
- 30. Lax I, et al. (1991) Noncontiguous regions in the extracellular domain of EGF receptor define ligand-binding specificity. Cell Regul 2(5):337-345.
- 31. Gadella TW, Jr, Jovin TM (1995) Oligomerization of epidermal growth factor receptors on A431 cells studied by time-resolved fluorescence imaging microscopy. A stereochemical model for tyrosine kinase receptor activation. J Cell Biol 129(6):1543-1558.
- 32. Berkers JA, van Bergen en Henegouwen PM, Boonstra J (1991) Three classes of epidermal growth factor receptors on HeLa cells. J Biol Chem 266(2):922-927.
- 33. Fleishman SJ, Schlessinger J, Ben-Tal N (2002) A putative molecular-activation switch in the transmembrane domain of erbB2. Proc Natl Acad Sci USA 99(25):15937-15940.
- Lemmon MA, Treutlein HR, Adams PD, Brünger AT, Engelman DM (1994) A dimerization motif for transmembrane alpha-helices. Nat Struct Biol 1(3):157-163.
- 35. Moriki T, Maruyama H, Maruyama IN (2001) Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain. J Mol Biol 311(5):1011-1026.
- 36. Mineev KS, et al. (2010) Spatial structure of the transmembrane domain heterodimer of ErbB1 and ErbB2 receptor tyrosine kinases. J Mol Biol 400(2):231-243
- 37. Prakash A, Janosi L, Doxastakis M (2011) GxxxG motifs, phenylalanine, and cholesterol guide the self-association of transmembrane domains of ErbB2 receptors. Biophys J 101(8):1949-1958.
- 38. Prakash A, Janosi L, Doxastakis M (2010) Self-association of models of transmembrane domains of ErbB receptors in a lipid bilayer. Biophys J 99(11):3657-3665.
- 39. Jura N, et al. (2011) Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. Mol Cell 42(1):9-22.
- 40. Red Brewer M, et al. (2009) The juxtamembrane region of the EGF receptor functions as an activation domain. Mol Cell 34(6):641-651.
- 41. Choowongkomon K, Carlin CR, Sönnichsen FD (2005) A structural model for the membrane-bound form of the juxtamembrane domain of the epidermal growth factor receptor. J Biol Chem 280(25):24043-24052.
- 42. Li S, Kussie P, Ferguson KM (2008) Structural basis for EGF receptor inhibition by the therapeutic antibody IMC-11F8. Structure 16(2):216-227.
- 43. Ferguson KM, et al. (2003) EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol Cell 11(2):507-517.
- Lax I, et al. (1989) Functional analysis of the ligand binding site of EGF-receptor utilizing chimeric chicken/human receptor molecules. EMBO J 8(2):421-427.
- 45. Kohda D. et al. (1993) A 40-kDa epidermal growth factor/transforming growth factor alpha-binding domain produced by limited proteolysis of the extracellular domain of the epidermal growth factor receptor. J Biol Chem 268(3):1976-1981.
- 46. Schmitz KR. Bagchi A. Roovers RC. van Bergen en Henegouwen PM. Ferguson KM (2013) Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. Structure 21(7):1214-1224.
- Miljan EA, et al. (2002) Interaction of the extracellular domain of the epidermal growth factor receptor with gangliosides. J Biol Chem 277(12):10108-10113.
- 48. Lopez PH, Schnaar RL (2009) Gangliosides in cell recognition and membrane protein regulation. Curr Opin Struct Biol 19(5):549-557.
- 49. Arkhipov A, Shan Y, Kim ET, Shaw DE (2014) Membrane interaction of bound ligands contributes to the negative binding cooperativity of the EGF receptor. PLOS Comput Biol 10(7):e1003742.
- 50. Mi LZ, et al. (2011) Simultaneous visualization of the extracellular and cytoplasmic domains of the epidermal growth factor receptor. Nat Struct Mol Biol 18(9):984-989.
- 51. Jura N, et al. (2009) Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. Cell 137(7):1293-1307.
- 52. van Meer G, Voelker DR, Feigenson GW (2008) Membrane lipids: Where they are and how they behave. Nat Rev Mol Cell Biol 9(2):112-124.
- 53. Czogalla A, Grzybek M, Jones W, Coskun U (2014) Validity and applicability of membrane model systems for studying interactions of peripheral membrane proteins with lipids. Biochim Biophys Acta 1841(8):1049-1059.
- Saffarian S, Li Y, Elson EL, Pike LJ (2007) Oligomerization of the EGF receptor investigated by live cell fluorescence intensity distribution analysis. Biophys J 93(3):1021-1031.
- 55. Kaiser HJ, et al. (2011) Lateral sorting in model membranes by cholesterol-mediated hydrophobic matching. Proc Natl Acad Sci USA 108(40):16628-16633.
- 56. McLaughlin S, Smith SO, Hayman MJ, Murray D (2005) An electrostatic engine model for autoinhibition and activation of the epidermal growth factor receptor (EGFR/ ErbB) family. J Gen Physiol 126(1):41-53.
- 57. Sengupta P, et al. (2009) EGFR juxtamembrane domain, membranes, and calmodulin: Kinetics of their interaction. Biophys J 96(12):4887-4895.