J Mol Model (2013) 19:931–941
DOI 10.1007/s00894-012-1613-y

ORIGINAL PAPER

Simulation of homology models for the extracellular domains (ECD) of ErbB3, ErbB4 and the ErbB2–ErbB3 complex in their active conformations

Juan Felipe Franco-Gonzalez · Javier Ramos · Victor L. Cruz · Javier Martínez-Salazar

Received: 11 May 2012 / Accepted: 26 September 2012 / Published online: 23 October 2012
© Springer-Verlag Berlin Heidelberg 2012

Abstract Epidermal growth factor receptors (EGFR) are associated with a number of biological processes and are becoming increasingly recognized as important therapeutic targets against cancer. In this work, we provide models based on homology for the extracellular domains (ECD) of ErbB3 and ErbB4 in their active conformations, including a Heregulin ligand, followed by further refinement of the models by molecular dynamics simulations at atomistic scale. We compare the results with a model built for ErbB2 based on crystallographic information and analyze the common features observed among members of the family, namely, the periscope movement of the dimerization arm and the hinge displacement of domain IV. Finally, we refine a model for the interaction of the ECDs corresponding to a ErbB2–ErbB3 heterodimer, which is widely recognized to have a high impact in cancer development.

Keywords EGFR receptors · Homology modeling · Molecular dynamics simulation

Introduction

Epidermal growth factor receptors (EGFR) control multiple cellular processes, including cellular proliferation, survival, differentiation and migration [1]. The EGFR family consists of four different members: EGFR/ErbB1, ErbB2, ErbB3 and ErbB4. These receptors are formed by three differentiated regions: an extracellular ligand-binding domain (ECD) containing four sub-domains (I–IV), a single transmembrane α-helix (TMH) spanning the cellular membrane and intracellular juxtamembrane, tyrosine-kinase and autophosphorylation domains. It is well established that epidermal growth factors (EGF) bind to the ECD domain, promoting EGFR dimerization and increasing the tyrosine-kinase activity of its intracellular domain.

The orientation of the four ECD sub-domains has been revealed by recent crystallographic studies. On the basis of these studies, tethered (ErbB3 [2], ErbB1/EGF [3], ErbB1/Cetuximab [4],ErbB4 [5]) and extended (EGF-bound [6] and TGFα-bound [7] truncated ErbB1 dimers, trastuzumab [8] and pertuzumab [9] ErbB2 complexes and full EGF-bound ErBB1 dimer [10]) conformations have been proposed for ECD-EGFR receptors. The tethered or autoinhibited conformation ties domains II and IV, forming intramolecular interactions. Accordingly, the dimerization arm in domain II is buried by domain IV, impairing the formation of EGFR dimers. On the contrary, interactions between domains II and IV are broken in the active untethered or extended conformation, provoking a release of the dimerization arm, which can now dimerize with other EGFR monomers.

The ECD of ErbB1, ErbB3, ErbB4 contains a specific region that binds EGF. This binding is thought to promote intracellular signaling cascades that regulate cellular growth and proliferation. Ligand binding promotes signaling activation by homo- or hetero-dimerization in the membrane surface, changing the receptor conformation from a tethered structure to an extended dimerization-exposed arm conformation [11]. Other authors, however, claim that the ligand can alter the equilibrium between previously formed active/

Electronic supplementary material The online version of this article (doi:10.1007/s00894-012-1613-y) contains supplementary material, which is available to authorized users.

J. F. Franco-Gonzalez · J. Ramos · V. L. Cruz (✉) · J. Martínez-Salazar
BIOPHYM, Macromolecular Physics Department,
Instituto de Estructura de la Materia, CSIC,
Serrano 113 bis,
28006 Madrid, Spain
e-mail: victor.cruz@iem.cfmac.csic.es

inactive dimers [11, 12]. Unlike other members of the family, the ErbB2 receptor can adopt a fixed conformation (extended conformation) resembling the ligand-activated state, which is able to dimerize even in the absence of ligand [13]. These receptors are engaged in the regulation of many processes such as cell proliferation, differentiation and apoptosis. The abnormal regulation of these receptors generates a number of human diseases, such as cancer [1, 14].

As mentioned above, crystal structures for the extended configurations of ErbB1 [6, 7] and ErbB2 [8, 9] are available. Additionally, homology structures for ErbB3 and ErbB4 have been built using as template the crystal structure of 2:2 EGF:EGFR complex (PDB code: 1IVO) [15]. Note that this template lacks the disordered domain IV, and homology structures can be obtained without this critical domain. It is only recently that the crystal structure of the 2:2 EFG:EGFR dimer containing domain IV has been reported [10]. Therefore, we propose to use this structure as a template to model a more complete structure of the ErbB3 and ErbB4 receptors.

The biological activity of these receptors is known to depend on the formation of homo or heterodimers in such a way that the intracellular kinase domains in close contact can interact, initiating a signaling cascade that ends with cellular proliferation [16]. The complex formed by ErbB2 and ErbB3 has been recognized as being of pivotal importance in cancer development and evolution [17].

Considering some theoretical work concerning ECDs, Fuentes et al. [18] have carried out several molecular dynamics (MD) studies on ErbB2 along with two different antibodies, trastuzumab (Tzb) and pertuzumab. They estimated the binding free energy for several receptor/antibody complexes. They reported that trastuzumab has a higher affinity for apo ErbB2 than pertuzumab. Furthermore, the epitope for trastuzumab is domain IV whereas pertuzumab is bound to the dimerization arm located in domain II. Subsequently, they found an increase in affinity when both antibodies are bound to the receptor.

On the other hand, Du et al. [19] performed microsecond MD simulations of ErbB4 tethered with its endogenous ligand neuregulin1β (NRG1β). While the conformational transition of the ECR-ErbB4/NRG1β complex from a tethered inactive conformation to an extended active-like form is observed clearly in the simulation, the conformational change of ECR-ErbB4 is not. Therefore, it could be proposed that ligand binding is indeed the active driving force for the conformational transition and further dimerization to occur. These authors constructed an energy landscape for the conformational transition of ECR-ErbB4/NRG1β complex and reported that the energy barrier for the tether opening has a value of 2.7 kcal/mol, which is close to the experimental value (1–2 kcal/mol) reported for ErbB1.

Additionally, Bagossi et al. [20] reported the first ErbB2/Tzb:ErbB2/Tzb complex from homology based on the X-ray or nuclear magnetic resonance structures of extracellular, transmembrane, and intracellular domains. They predicted some favourable dimerization interactions for the extracellular, transmembrane, and protein kinase domains in the model of a nearly full-length dimer of ErbB2, which may act in a coordinated fashion in ErbB2 homodimerization.

The objective of the present study was to build homology models of the ECDs corresponding to ErbB3 and ErbB4 receptors in their extended conformations along with a model of the back-to-back ErbB2/ErbB3:HRG-α heterodimer.

Water-equilibrated models of the ErbB2, *HRG-α*:ErbB3 and *HRG-α*:ErbB4 complexes were built using as templates the X-ray crystal structure of the ErbB2/Tzb (PDB code: 1NZ8) and 2:2 EFG:EFGR dimer (PDB code: 3NJP) complexes, respectively, as starting structures, due to the high homology of this receptor family. MD simulations (100 ns) were performed and details of the interactions and intrinsic motions were investigated for each model through principal component analysis (PCA). Additionally, a homology structure of the ECDs of the back-to-back ErbB2:ErbB3: *HRG-α* complex were built and refined through multi-nanosecond MD simulation. To our best of our knowledge, this is the first computational model of this important complex.

Computational methods

Homology models for the ECD of HRG-α:ErbB3, HRG-α:ErbB4 and back-to-back ErbB2:ErbB3:HRG-α complexes

The amino acid sequences of HRG-α, ErbB3 and ErbB4 were obtained from the SwissProt database (P04626, P21860,Q15303). Homology models of the HRG-α:ErbB3 and HRG-α:ErbB4 complexes were built using as a template the X-ray crystal structure of chain A of the 2:2 EFG:EGFR dimer (PDB code: 3NJP) [10]. Unlike the published homology models for HRG-α:ErbB3 and HRG-α:ErbB4 based on the 1IVO structure [15], our template allows the region of domain IV to be included in the homology model. The importance of the orientation of this tight domain has been recognized because of the dependence of intracellular protein kinase domain orientation on domain IV disposition [6, 7].

All models were generated using the PRIME application of the Schrödinger Suite 2011 (Schrödinger, New York, NY). First, pairwise sequence alignments between ErbB3 and ErbB4 and the EFGR template sequences were performed. The BLAST alignment gives 61 % and 63 % of positives (percentage of residues that are positive matches to the BLOSUM62 similarity matrix [21]) and 46 % and 47 % of identities (percentage of residues that are identical

between sequences) using 1 % and 1 % of gaps between the template and ErbB3 and ErbB4, respectively. However, this simple alignment does not assign adequately the secondary structure between the template and the query, as predicted by SSpro program [22] . For this reason, we adopted an alignment algorithm (single template algorithm, STA) that takes into account both secondary structure matching and profile-sequence matching. This procedure minimizes the inaccuracy in a single secondary structure prediction at the expense of increasing the percentage of gaps. This algorithm was designed specially for protein sequences with medium-to-high sequence identity (>25 %). The 3D-structures were built by replicating the backbone atom coordinates for the aligned regions and side chains of conserved residues, followed by optimization of the side chains and non-template residues. Finally, the gaps were filled by insertions and final closing of deletions in the alignment. In all cases, the gap length was less than 15. Finally, the homology model coordinates were minimized with Schrödinger [23] using OPLS-2005 with the default threshold of 0.05 kJ/mol as the convergence criteria.

The same procedure was used for the homology model of the HRG-α peptide ligand taking as template the EGF peptide of the 3NJP structure. The final HRG-α:ErbB3 and HRG-α:ErbB4 models were built using both sets of coordinates, i.e., receptor and ligand moieties, respectively.

Model of the ErbB2 receptor

The initial model of the ErbB2 ectodomain was extracted directly from the 3-D crystal structure deposited with the Protein Data Bank server (PDB code: 1N8Z) [8]. The missing residues N^{124}–A^{132} (domain I), E^{325}–G^{327} (domain II), G^{383}–G^{386} (domain III) and G^{603}–P^{612} (domain IV) were modeled based on homologous sequences using the PRODAT database implemented in Sybyl 8.0 [24]. The loop fragment that affords the best geometric fit as monitored by the homology score and RMS fit, was incorporated automatically into the model [24, 25] . The side chains were built using the rotamer library of Sybyl by taking a scan angle of 30° and a VDW factor of 0.9. The selected side-chain conformation is the one that presents the fewest bumps on the rest of the molecule. Finally, the structure was relaxed over 2,500 steps using the steepest descent minimization algorithm as implemented in GROMACS 4.5.3 [26]

We assume that the pKa of the individual amino acid residues at physiological pH does not change when they are assembled into the protein receptor. Thus, histidine (H) residues were kept neutral while lysine (L) and arginine (R) were protonated and aspartic (D) and glutamic (E) acids were deprotonated.

The resulting total charge for the complex was −10 e. The system was then solvated by adding 60,717 water molecules and 10 Na^+ ions to yield an electrically neutral system. The system was finally equilibrated in a 2 ns NPT-MD simulation with position restraint for all protein atoms.

Molecular dynamics

The OPLS force field [27–29] for protein and SPC model [30] for water were used throughout this work. Short-range repulsion-dispersion interactions were truncated smoothly at 10 Å. The particle mesh Ewald (PME) method [31, 32] was used to calculate long-range electrostatic interactions, by means of a maximum grid spacing of 2.5 Å and using fourth-order (cubic) interpolation for the fast Fourier transforms. The temperature was kept constant at 300 K by coupling the protein, the ions and the solvent independently to an external bath using the Berendsen algorithm [33] with a coupling constant of 0.2 ps.

We used isotropic scaling for the pressure (1 bar) and a coupling constant of 1.0 ps and a compressibility of $4.5\times 10^{-5}\,bar^{-1}$ when applying the Berendsen algorithm [33]. The dynamics were integrated using the velocity Verlet integrator, with a time step of 2 fs and bonds constrained using the LINCS algorithm [34].

Production dynamics was performed at constant pressure and temperature (NPT ensemble) releasing all constraints on the heavy atoms during 100 ns, and storing the trajectory every 10 ps. All minimizations, restrained and unrestrained MD runs were performed with GROMACS 4.5.3 [26]. Molecular graphics were drawn using the VMD 1.8.7 package [35].

PCA, hydrogen bonds and contact maps

PCA is a method that takes the trajectory of long MD simulations and calculates the dominant modes in the motion of the molecule. Thus, the conformational space is reduced, resulting in few relevant collective degrees of freedom over which long-range fluctuation can be studied [36, 37]. PCA diagonalizes the covariance matrix of the atom fluctuations from their average trajectory. In this framework, the larger eigenvalues capture the larger fluctuations fraction. The ordering of these eigenvalues gives rise to a small set of modes that capture most of the protein's fluctuation. We performed PCA in order to identify the most relevant motions occurring in the EGFR family. In this work, we make use of the first five eigenvectors, which were projected along the MD trajectory.

Hydrogen bonds (HB) are considered to exist when both distance between the donor (D) and the acceptor (A) is less than 0.30 nm and the DHA (hydrogen-donor-acceptor) angle is less than 30°.

The contact maps show the smallest distance between any pair of atoms belonging to two different residues. The output is a symmetrical matrix of smallest distances between all residues. Plotting these matrices for

different time-frames is a useful tool with which to analyze changes in the structure, HB networks and hydrophobic contacts.

The root-mean square fluctuation (RMSF) for each residue was calculated using the g_rmsf tool from GROMACS.

The change in secondary structure elements during the simulation was monitored using the program DSSP (Define Secondary Structure of Proteins) [38].

Results and discussion

Comparative homology modeling

Three-dimensional homology models were built for the ECD of HRGα:ErbB3 and HRGα:ErbB4 complexes, using as template the X-ray structure corresponding to chain A of the 2:2 EGF:EGFR dimer (PDB code: 3NJP) [10]. Full alignments of the models are shown in Supplementary information. The EGFR template has a high sequence similarity of 41 % and 42 % for the ErbB3 and ErbB4 proteins, respectively. Cysteine residues forming disulfide bonds in both receptor and ligand were well conserved during the alignment process. Figure 1 shows the Cα backbone superposition of the refined HRGα:ErbB3 and HRGα:ErbB4 on the X-ray structure of the EGF:EGFR. The root mean squared deviations (RMSD) are 0.8 and 0.7 Å for both ErbB3 and ErbB4 Cα backbones, respectively. As expected, a high degree of structural similarity was observed due to the high sequence homology. Furthermore, all secondary motifs are well captured as seen in Fig. 1.

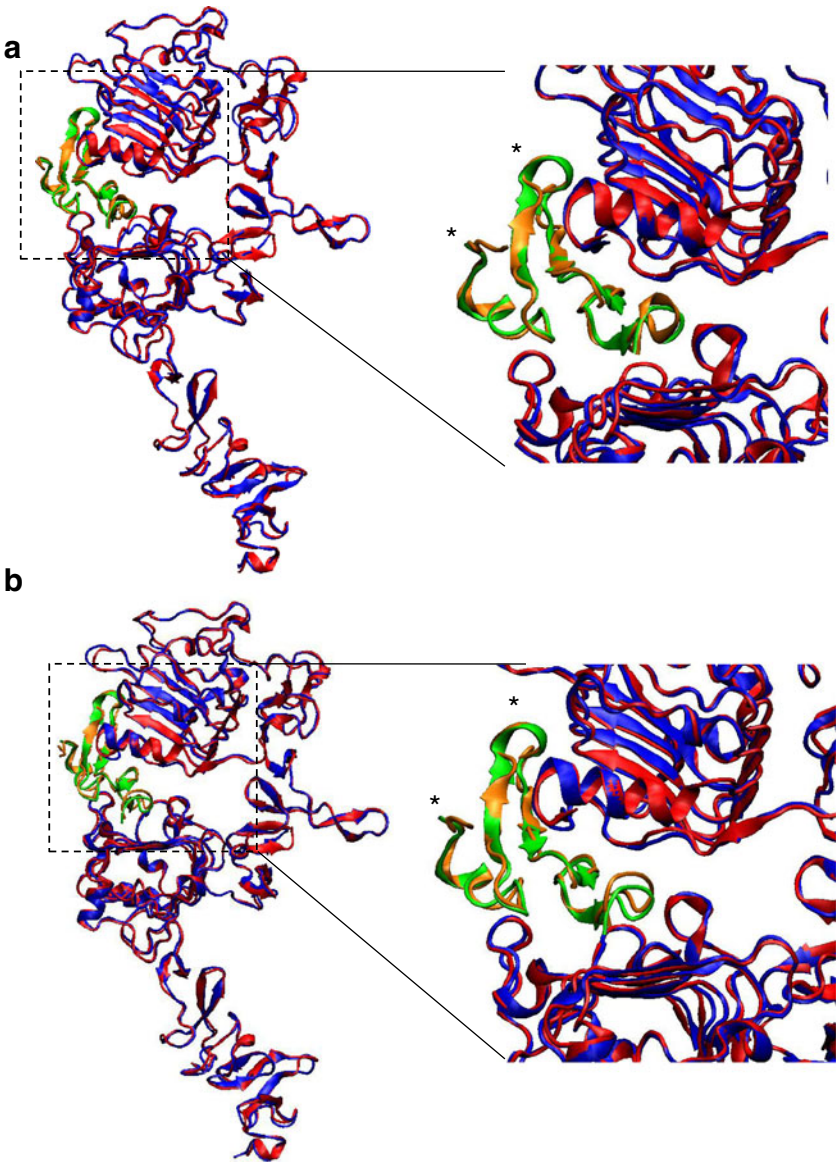

Fig. 1 Refined homology model (**a**) HRGα(*orange*):ErbB3(*blue*) and (**b**) HRGα(*orange*):ErbB4(*blue*) superimposed on the 3NJP PDB-code X-ray structure EGF(*green*):EGFR(*red*) (PDB code 3NJP). The areas on the *right* are enlargements showing ligand–receptor interactions. *Asterisks* mark the ligand loop with largest deviations between the target and the homolog (see text)

Table 1 Ramachandran plot of HRGα:ErbB3, HRGα:ErbB4 and ErbB2:ErbB4:HRGα homology models calculated with the PROCHECK software. The 3NJP and 1N8Z x-ray structures are shown as reference. Glycine and proline residues are not taken into account

	EGF:EGFR[a]	ErbB2[b]	HRGα:ErbB3	HRGα:ErbB4	ErbB2/ErbB3:HRGα
% in most favored regions	78.2	82.6	72.2	72.5	72.3
% in additional allowed regions	19.9	15.3	23.9	24.8	23.5
% in generously allowed regions	1.1	2.0	2.7	1.2	2.7
% in disallowed regions	0.8	0.0	1.3	1.5	1.5

[a] Chain A of the X-ray structure reported in the 3NJP pdb file

[b] Chain C of the X-ray structure reported in the 1N8Z pdb file

The HRG-α ligand has been modeled using the EGF structure in the 3NJP complex as template for both ErbB3 and ErbB4 cases. The sequence identity of the EGF template and HRGα ligand is found to be around 25 %. There are a couple of NMR-solved structures for HRG-α (PDB code: 1HAF [39] and 1HRF [40]) . However, the EGF ligand in the 3NJP complex is preferred here over the NMR structure in solution, because in principle the ligand–protein interactions are better modeled in the 3NJP structure where the surroundings are taken into account. RMSD values of 3.1 and 3.0 Å were calculated between the EGF Cα backbone atoms in 3NJP and HRGα in ErbB3 and in ErbB4 homology models, respectively. These high values are attributable to different residues in the region of 1–6 (unbounded loop) and 23–28 (flexible loop between β-sheet structures) of EGF and HRG-α ligands (asterisk in the zoomed area of Fig. 1).

The models were evaluated using Ramachandran plots calculated by the PROCHECK program. The percentage occupancies for favored, allowed, generously allowed and disallowed regions are collected in Table 1. A percentage of 96.1 and 97.3 % of the residues are in favored and allowed regions for the HRG-α:ErbB3 and HRG-α:ErbB4 models, respectively. After these analyses, the quality of the HRGα:ErbB3 and HRGα:ErbB4 complexes seem to be good enough for further study.

The homology model of the ErbB2:ErbB3:HRGα extracellular complex fitted on the X-ray structure of EGFR:EGF homodimer is shown in Fig. 2. The RMSD of the Cα backbone atoms are 3.1, 0.8 and 3.0 Å for ErbB2, ErbB3 and HRGα ligand, respectively. The flexible loop made up of the 70–114 residues in ErbB2 is causing the high RMSD value (marked with an asterisk in Fig. 2). This value is reduced to 0.7 Å once the loop is discarded. Again, the high values of RMSD for the ligand are attributable to the flexible loops discussed above. Following the Ramachandran plot, the 95.8 % of the residues are in favored and allowed regions for the models. Thus, the ErbB2:ErbB3:HRGα extracellular complex is good enough for further MD studies.

Molecular dynamics analysis

We have performed MD simulations of each extracellular system in Table 1 (except for the EGF:EGFR complex) for at least 100-ns. The aim of these simulations was to study the stability and the dynamics of the homology models.

Atomistic dynamics of the ErbB2, HRGα:ErbB3 and HRGα:ErbB4 molecules

First, the time evolution of different domain RMSD (Fig. 3a) was used to track conformational motions along the trajectory. No global conformational switch from the extended to the tethered structure was detected in any case. Thus, the extended conformations remain stable for ErbB2, HRGα:ErbB3 and HRGα:ErbB4 systems as corresponding

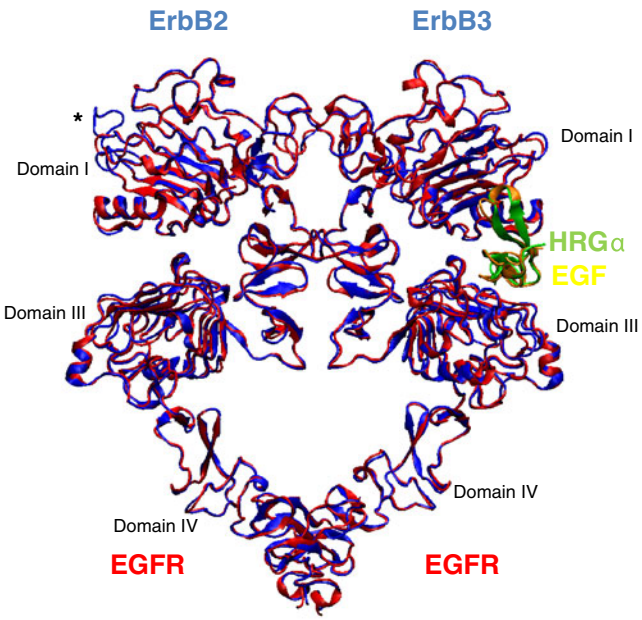

Fig. 2 ErbB2:ErbB3:HRGα extracellular heterodimer complex (*blue* for ErbB2 and ErbB3 receptors and *green* for HRG α ligand) using the (EGFR:EGF)2 homodimer (*red* and *yellow*) as template (PDB code 3NJP). The *asterisk* marks the residues from 70 to 114 (flexible loop) where the major differences between the target and the homology model are found

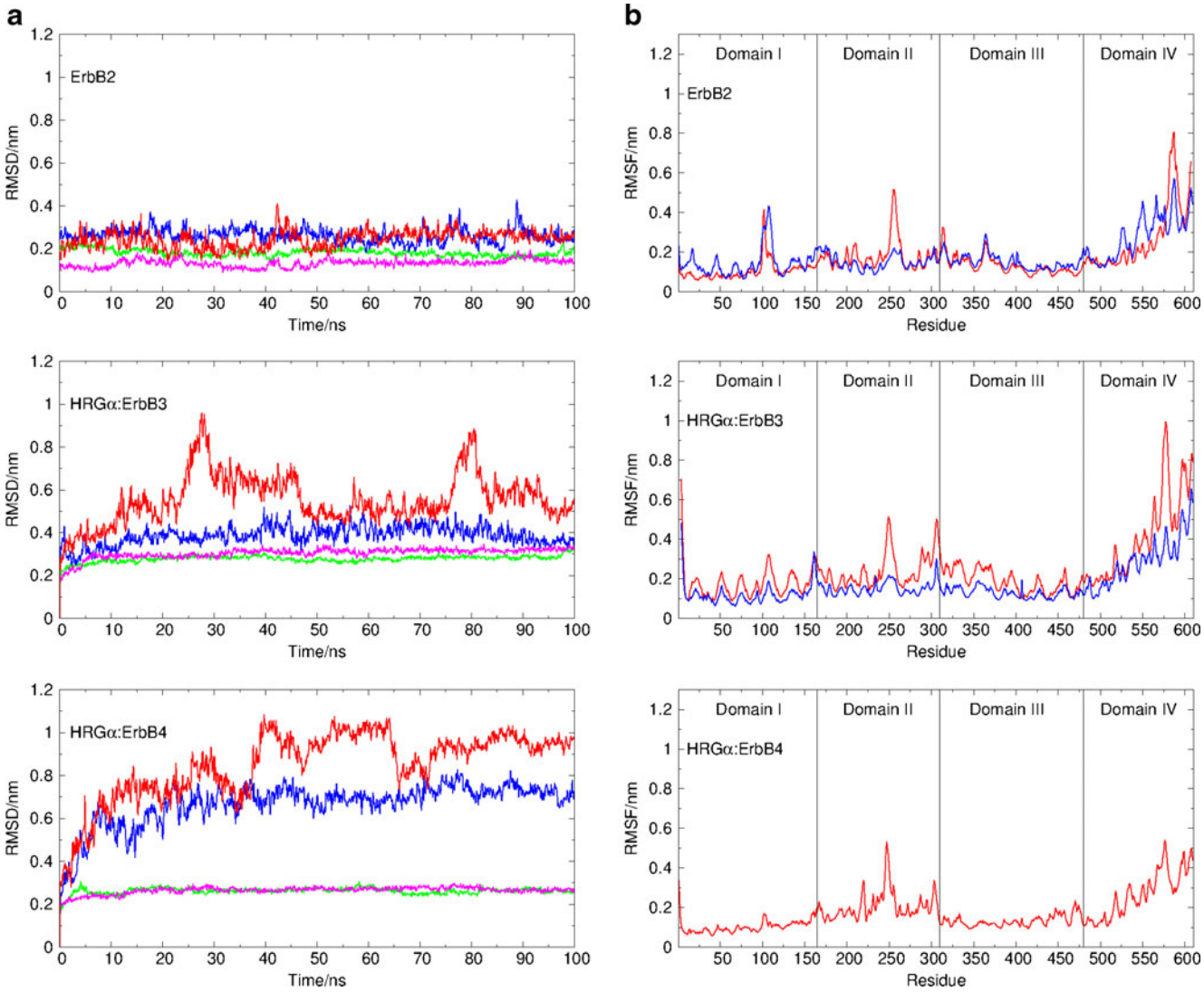

Fig. 3 **a** Root mean square deviation (RMSD) of the different domains as a function of time for ErbB2 (*top*), HRGα:ErbB3 (*middle*) and HRGα:ErbB4 (*bottom*) systems. Domains I, II, III and IV are shown as *green*, *blue*, *magenta* and *red lines*, respectively. **b** Root-mean square fluctuation (RMSF) as a function of the residue number in backbone atoms. The monomers ErbB2 (*top*), HRGα:ErbB3 (*middle*) and HRGα:ErbB4 (*bottom*) are shown as *red lines*. The *blue lines* corresponds to ErbB2 and HRGα:ErbB3 in the heterodimer complex. Roman numbers are used to number the most important peaks. The last 50 ns of the trajectory were taken in all cases

to active states. Domains I and III remain stable along the whole trajectory to values smaller than 2 Å from the crystal structure (ErbB2) or homology models (HRGα:ErbB3 and HRGα:ErbB4). Domains II and IV for ErbB2 are stable with respect to the reference crystal structure, showing only a small deviation of around 3 Å. On the contrary, domains II and IV for the HRGα:ErbB3 and HRGα:ErbB4 receptors exhibit deviations in the range of 4–7 and 6–10 Å, respectively. The RMSD profiles for HRGα:ErbB3 and HRGα:ErbB4 complexes are similar, but the domain II deviations, which are larger for HRGα:ErbB4 (7 Å versus 4–5 Å, blue lines in Fig. 3a), point to a more movable dimerization arm. In summary, the largest deviations were observed for homology-based structures along the MD trajectories.

Additionally, some remarkable residue moves were observed through RMSF analysis. As shown in Fig. 3, the intense peak observed in domain I (peak I) for ErbB2 corresponds to the motion of a flexible loop (N^{124}–A^{132}). On the contrary, this loop becomes more constrained in ErbB3 and ErbB4 proteins as a consequence of the presence of the HRGα ligand. It should be noted that this loop has not been solved experimentally for the ErbB2 extracellular domain by X-ray techniques, suggesting that the loop is very flexible in that protein [8]. The dimerization arm residues exhibit high RMSF values for all proteins (peak II in Fig. 3), showing a periscope-like motion [18]. This motion can assist the formation of dimers between different members of the EGFR family. Finally, domain IV is very mobile with the highest RMSF values (peak III) in all

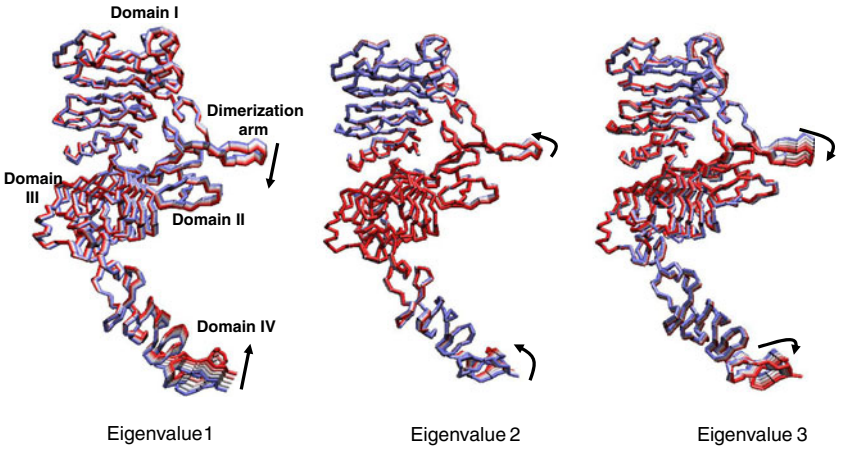

Fig. 4 Dominant motions represented by the first three principal component analysis (PCA) eigenvectors in EGFR systems. Several projections are shown in different colors between the two extreme points(*blue* and *red colors*). *Arrows* indicate the direction of motion along the molecular dynamics (MD) simulation for each eigenvector. This figure shows the eigenvectors on the ErbB2 molecule. Eigenvectors for the HRGα:ErbB3 and HRGα:ErbB4 systems are similar, so are shown in separate Figures in the Supplementary information section

cases for the proteins alone, corresponding to a hinge motion within this domain. This flexibility could explain the difficulty encountered by other studies in elucidating domain IV by crystallographic techniques on the extended conformations of protein monomers [7, 8]. Although domain II is very flexible in all members of the EGFR family, inspection of both trajectory and RMSD values suggests that ErbB4 is the most flexible.

Let us focus now our analysis on the slowest modes that could provide information on the collective motions of EGFR domains. PCA was performed on the Cα atoms for the three molecules. Despite the fact that PCA was calculated for a single trajectory, some useful qualitative information can be extracted from this analysis. The first three eigenvectors account for ca. 80 % of the motion in ErbB2, HRGα:ErbB3 and

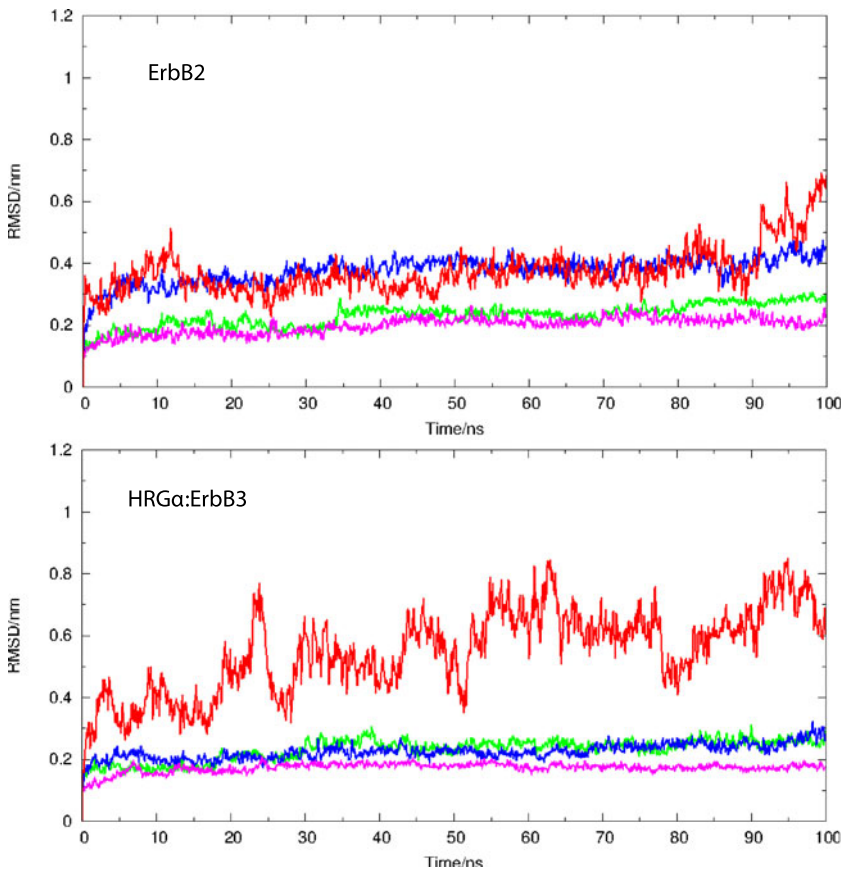

Fig. 5 RMSD of the domains as a function of time for ErbB2 (*top*) and HRGα:ErbB3 (*bottom*) systems in the heterodimer complex. Domains I, II, III and IV are shown as *green*, *blue*, *magenta* and *red lines*, respectively

HRGα:ErbB4 molecules, respectively. The principal motions for the first three eigenvectors of the ErbB2 molecule are shown in Fig. 4 (for ErbB3 and ErbB4 the motions are shown as Supplementary information because they are similar. The corresponding videos are available as Supplementary information). A concerted hinge movement of domain IV is observed clearly for the first PCA eigenvector. The other two eigenvectors show clockwise and counterclockwise torsions of domains II and IV, respectively. Thus, PCA analysis confirms that domains II and IV are very flexible, in agreement with the previous discussion.

Summarizing, MD analysis of the free EGFRs led to the conclusion that great flexibility of domains II and IV exists in the extended structures. It could be hypothesized that this intrinsic flexibility would increase the possibility of forming dimers among different members of the EGFR family.

Atomistic dynamics of the back-to-back ErbB2:ErbB3:HRGα complex

Global conformation dynamics The homology model for the *ErbB2:ErbB3:HRGα* heterodimer is shown in Fig. 2. We used this model as the initial structure for a 100 ns MD simulation. The overall change in conformation of the heterodimer from the homology structure can be tracked by plotting the RMSD of each domain (Fig. 5). All domains in the monomers remain stable within less than 0.4 nm from the homology model and they all relax over the first 10 ns. The exception is domain II in ErbB3, which has higher RMSD values (ca. 0.6 nm) and large fluctuations, even at the end of the simulation. Moreover, the flexibility of each domain of the heterodimer was quantified by calculating the RMSF for Cα atoms in each residue (Fig. 3b). The fluctuations in domains II and IV (peaks II and III) are reduced in the heterodimer (blue lines in Fig. 3b) as compared to the free monomers (red lines in Fig. 3b). Therefore, as expected, interactions between the dimerization arms (domain II) and domain IV of both ErbB2 and ErbB3 monomers reduce the flexibility of such domains in the heterodimer.

Intermolecular dimerization interfaces Experimentally, the crystal structure of the $(EGFR:EGF)_2$ homodimer is mediated by intermolecular interactions involving both the dimerization arm in domain II [6, 7, 10] and the C terminus located at domain IV [10]. In the present case, the dimerization arm in domain II of any monomer (i.e., ErbB2) tends to interact with a region between domains I and II of the neighboring receptor. Interactions between residues placed at domain IV are also observed (see Fig. 2). The most important non-bonded close contacts between the ErbB2 and ErbB3 are collected in Table 2. These contacts take into account salt bridges, HBs, aromatic interactions and hydrophobic contacts. As it can be seen in Table 2, the domain II–domain II interface is made up of both hydrophobic and HB interactions in which the 15-residue dimerization arms go through the pocket region of domains I and II of the neighboring receptor. However, a persistent interaction with domain III of ErbB2 was observed only for the ErbB3 dimerization arm. The domain IV–domain IV interfaces are shown in Fig. 6. As can be seen, two regions are kept from the last 50 ns of the trajectory. The first region (Fig. 6b) corresponds to residues A^{599}–L^{609} and S^{587}–E^{596} in ErbB2 and ErbB3, respectively. This region has a marked hydrophobic character with residues P^{602} and V^{605} in the ErbB2 receptor and P^{590}, V^{593} and L^{594} in the ErbB3 receptor. On the other hand, the second region (Fig. 6c) corresponds to residues C^{631}–C^{635} and C^{617}–G^{623} in ErbB2 and ErbB3,

Table 2 Interface non-bonded close contacts (salt bridges, H-bonds, aromatic and hydrophobic interactions) in the ErbB2:ErbB3:HRGα complex. The domain to which each amino acid belongs is given in parentheses. Only residue–residue contacts with averaged contact area above 20 $Å^2$ are displayed. The H-bonds reported fulfill the criteria that the donor–acceptor distance and angle cutoffs of 3.5 Å and 30° are maintained during at least 50 % of the last 50 ns of the trajectory. The distance in salt bridges is less than 4.0 Å

	ErbB2	ErbB3	Interaction type
Dimerization ErbB2 arm[a]	V^{273}(II)	R^{258}(II)	Hydrophobic
	Y^{275}(II)	C^{301}(II)	H-bond
	D^{278}(II)	T^{108}(I)	Hydrophobic
	F^{280}(II)	R^{267}(II), Y^{282}(II)	Hydrophobic, aromatic
	E^{281}(II)	A^{304}(II)	Hydrophobic
Dimerization ErbB3 arm[b]	T^{291}(II), C^{312}(III)	Y^{265}(II)	Hydrophobic, H-bonds
	H^{258}(II)	K^{267}(II)	Hydrophobic
	Q^{58}(I)	L^{268}(II)	H-bonds
	F^{292}(II), Y^{304}(II)	F^{270}(II)	Hydrophobic, aromatic
	L^{314}(III)	Q^{271}(II)	Hydrophobic
	T^{313}(III)	L^{272}(II)	Not assigned
Domain II - Domain II[c]	G^{224}(II)	N^{224}(II)	H-bond
	P^{233}(II)	Q^{213}(II)	H-bond
	K^{333}(II)	E^{321}(II)	Salt bridge
	R^{352}(II)	E^{273}(II)	Salt bridge
Domain IV - Domain IV[d]	P^{602} (IV), V^{605}(IV)	P^{590}(IV), V^{593}(IV), L^{594}(IV)	Hydrophobic
	H^{633}(IV)	C^{621}(IV)	H-bond
	A^{645}(IV)	L^{622}(IV)	H-bond

[a] Interaction between residues at the dimerization arm of the ErbB2 with the ErbB3 receptor. The arm dimerization in ErbB2 receptor is defined as residues 271–286

[b] Interaction between residues at the dimerization arm of the ErbB3 with the ErbB2 receptor. The arm dimerization in ErbB3 receptor is defined as residues 261–276

[c] Persistent hydrogen bonds between residues belonging to domain II

[d] Persistent contacts between residues located at domain IV

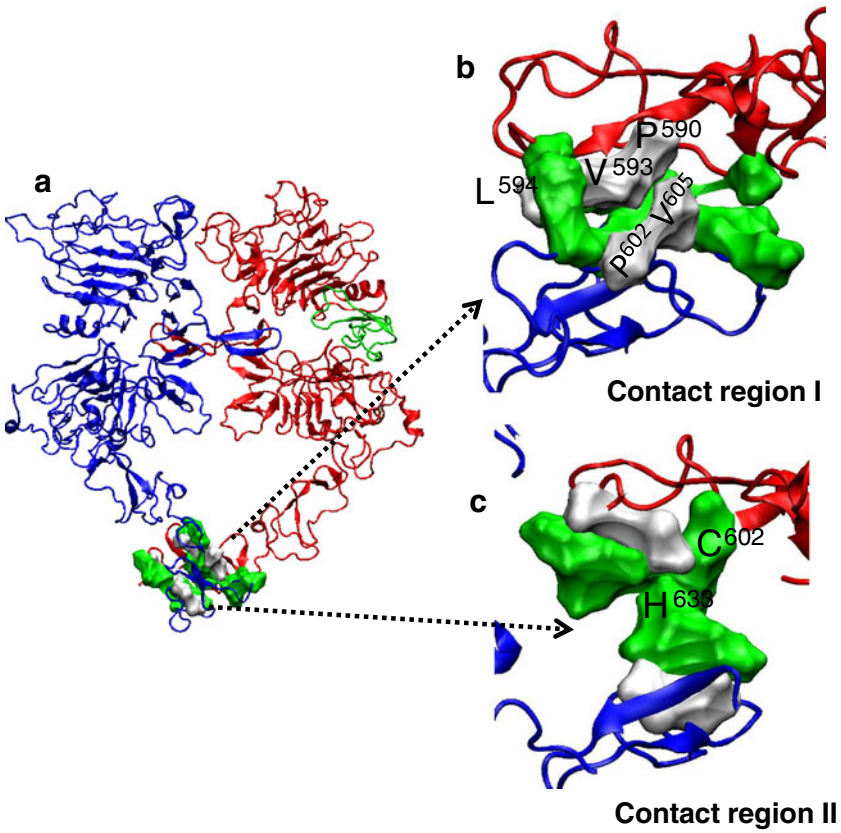

Fig. 6 **a** Domain IV–domain IV interface. Interactions in region defined as I (**b**) and II (**c**) are zoomed. *White surfaces* Hydrophobic contacts, *green surfaces* polar contacts. The main residues responsible for each interaction are labeled inside the figure. ErbB2 and ErbB3 receptors are shown as *blue* and *red* cartoon models

respectively. In this case, the interaction is maintained by the formation of two hydrogen bonds (see Table 2).

PCA was carried out on the MD trajectory to identify the most significant cooperative motions of the ErbB2:ErbB3:HRGα complex. The first three eigenvectors account for 75 % of the overall motions. These three eigenvectors consist mainly of concerted motions of both domains IV (not shown). Remarkably, the motion of domain II is more constrained in comparison with the ErbB2 and ErbB3 free receptors.

Conclusions

The present work provides a useful collection of homology models for the ECD ErbB receptors updated with the information provided by the latest crystallographic structures of suitable templates deposited with the Protein Data Bank. In particular, the model for ErbB3 and ErbB4 receptors includes the structure of domain IV which was absent in previous studies. The quality of the models is proved to be very satisfactory in view of the resulting RMSD differences with the template structures and Ramachandran maps showing suitable backbone torsion angle distributions.

These models were subsequently refined by MD simulation. It is worth mentioning that, according to the RMSD time evolution, the ErbB2 receptor, built from crystallographic data, exhibits the most stable structure along the whole simulation. A set of common features was found for all the receptors, namely a periscope movement of the dimerization arm in domain II, which confirms the findings of previous studies. What is more important is the remarkable flexibility found for domain IV. A hinge movement of this domain towards domains II and III was observed in all cases. In this context, PCA reveals that the first eigenvectors are associated to this collective movement.

We have also proposed a model for the interaction of ErbB2 and ErbB3. This complex forms one of the most biologically relevant heterodimers associated with aggressive carcinomas. To the best of our knowledge, the model proposed in our work is the first atomistic scale model for ECD interaction in this heterodimer. The structure presents the expected interaction between the two receptors through the dimerization arms in domain II, which immobilizes these domains with respect to the unbound structures. In addition to this, a weaker interaction through domain IV is also observed. However, the hinge movement observed in the separated receptors is also noticeable in the complex in an asymmetric way, being less mobile in ErbB2 than in ErbB3.

The biological consequences of this information are not so evident, and further studies need to be carried out.

However, the generation of the ECD ErbBs models presented in this work will serve as an starting point for a systematic study of these important receptors in order to provide clues for the development of more effective therapeutic strategies.

Acknowledgments Thanks are due to the Comision Interministerial de Ciencia y Tecnologia (CICYT) (MAT2009-12364 and MAT2012-36341 projects) for financial support. The authors also acknowledge "SGAI-CSIC" for technical support and allocation of computer time during the simulations. J.R. is grateful for financial support through the *Ramón y Cajal* program, contract RYC-2011-09585

References

1. Wieduwilt M, Moasser M (2008) The epidermal growth factor receptor family: biology driving targeted therapeutics. Cell Mol Life Sci 65(10):1566–1584. doi:10.1007/s00018-008-7440-8
2. Cho H-S, Leahy DJ (2002) Structure of the extracellular region of HER3 reveals an interdomain tether. Science 297(5585):1330–1333. doi:10.1126/science.1074611
3. Ferguson KM, Berger MB, Mendrola JM, Cho H-S, Leahy DJ, Lemmon MA (2003) EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol Cell 11(2):507–517
4. Li S, Schmitz KR, Jeffrey PD, Wiltzius JJW, Kussie P, Ferguson KM (2005) Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer Cell 7(4):301–311
5. Bouyain S, Longo PA, Li S, Ferguson KM, Leahy DJ (2005) The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. Proc Natl Acad Sci USA 102(42):15024–15029. doi:10.1073/pnas.0507591102
6. Ogiso H, Ishitani R, Nureki O, Fukai S, Yamanaka M, Kim J-H, Saito K, Sakamoto A, Inoue M, Shirouzu M, Yokoyama S (2002) Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell 110(6):775–787
7. Garrett TPJ, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO, Zhu H-J, Walker F, Frenkel MJ, Hoyne PA, Jorissen RN, Nice EC, Burgess AW, Ward CW (2002) Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor α. Cell 110(6):763–773
8. Cho H-S, Mason K, Ramyar KX, Stanley AM, Gabelli SB, Denney DW, Leahy DJ (2003) Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature 421 (6924):756–760 http://www.nature.com/nature/journal/v421/n6924/suppinfo/nature01392_S1.html
9. Franklin MC, Carey KD, Vajdos FF, Leahy DJ, de Vos AM, Sliwkowski MX (2004) Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. Cancer Cell 5 (4):317–328
10. Lu C, Mi n-Z, Grey MJ, Zhu J, Graef E, Yokoyama S, Springer TA (2010) Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal growth factor receptor. Mol Cell Biol 30(22):5432–5443. doi:10.1128/mcb.00742-10
11. Lemmon MA (2009) Ligand-induced ErbB receptor dimerization. Exp Cell Res 315(4):638–648
12. Chung I, Akita R, Vandlen R, Toomre D, Schlessinger J, Mellman I (2010) Spatial control of EGF receptor activation by reversible dimerization on living cells. Nature 464 (7289):783–787 http://www.nature.com/nature/journal/v464/n7289/suppinfo/nature08827_S1.html
13. Brennan PJ, Kumagai T, Berezov A, Murali R, Greene MI (2000) HER2/neu: mechanisms of dimerization/oligomerization. Oncogene 19(53):6093–6101. doi:10.1038/sj.onc.1203967
14. Yarden Y, Sliwkowski MX (2001) Untangling the ErbB signalling network. Nat Rev Mol Cell Biol 2(2):127–137. doi:10.1038/35052073
15. Luo C, Xu L, Zheng S, Luo X, Shen J, Jiang H, Liu X, Zhou M (2005) Computational analysis of molecular basis of 1:1 interactions of NRG-1beta wild-type and variants with ErbB3 and ErbB4. Proteins 59(4):742–756. doi:10.1002/prot.20443
16. Hynes NE, Lane HA (2005) ERBB receptors and cancer: the complexity of targeted inhibitors. Nat Rev Cancer 5(5):341–354. doi:10.1038/nrc1609
17. Baselga J, Swain SM (2009) Novel anticancer targets: revisiting ERBB2 and discovering ERBB3. Nat Rev Cancer 9(7):463–475. doi:10.1038/nrc2656
18. Fuentes G, Scaltriti M, Baselga J, Verma C (2011) Synergy between trastuzumab and pertuzumab for human epidermal growth factor 2 (Her2) from colocalization: an in silico based mechanism. Breast Cancer Res 13(3):R54
19. Du Y, Yang H, Xu Y, Cang X, Luo C, Mao Y, Wang Y, Qin G, Luo X, Jiang H (2012) Conformational transition and energy landscape of ErbB4 activated by neuregulin1β: one microsecond molecular dynamics simulations. J Am Chem Soc 134(15):6720–6731. doi:10.1021/ja211941d
20. Bagossi P, Horváth G, Vereb G, Szöllösi J, Tözsér J (2005) Molecular modeling of nearly full-length ErbB2 receptor. Biophys J 88(2):1354–1363
21. Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 89(22):10915–10919
22. Cheng J, Randall AZ, Sweredoski MJ, Baldi P (2005) SCRATCH: a protein structure and structural feature prediction server. Nucleic Acids Res 33(suppl 2):W72–W76. doi:10.1093/nar/gki396
23. (2011) Maestro, version 9.2, Schrödinger. LLC, New York, NY
24. SYBYL 8.0, Tripos International, St. Louis, MO
25. Dayhoff M, Schwartz R, Orcutt B (1978) Atlas of protein sequence and structure. National biomedical research foundation. National Biomedical Research Foundation, Washington DC, pp 5–345
26. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4(3):435–447. doi:10.1021/ct700301q
27. Jorgensen WL, Tirado-Rives J (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. J Am Chem Soc 110(6):1657–1666. doi:10.1021/ja00214a001
28. Jorgensen WL, Tirado-Rives J (2005) Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. Proc Natl Acad Sci USA 102(19):6665–6670. doi:10.1073/pnas.0408037102
29. Lawrence CP, Skinner JL (2003) Flexible TIP4P model for molecular dynamics simulation of liquid water. Chem Phys Lett 372(5–6):842–847
30. Berendsen H, Postma J, van Gunsteren W, Hermans J (1981) Interaction models for water in realtion of protein hydration. In: Pullman B (ed) Intermolecular forces. Reidel, Dordrecht, pp 331–342
31. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N [center-dot] log(N) method for Ewald sums in large systems. J Chem Phys 98(12):10089–10092
32. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103(19):8577–8593
33. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81(8):3684–3690

34. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: A linear constraint solver for molecular simulations. J Comput Chem 18(12):1463–1472. doi:10.1002/(sici)1096-987x(199709)18:12<1463::aid-jcc4>3.0.co;2-h
35. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14(1):33–38, 27–38
36. Amadei A, Linssen ABM, Berendsen HJC (1993) Essential dynamics of proteins. Proteins Struct Funct Bioinforma 17(4):412–425. doi:10.1002/prot.340170408
37. Amadei A, Linssen AB, de Groot BL, van Aalten DM, Berendsen HJ (1996) An efficient method for sampling the essential subspace of proteins. J Biomol Struct Dyn 13(4):615–625
38. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22(12):2577–2637. doi:10.1002/bip.360221211
39. Jacobsen NE, Abadi N, Sliwkowski MX, Reilly D, Skelton NJ, Fairbrother WJ (1996) High-resolution solution structure of the EGF-like domain of heregulin-α‡. Biochemistry 35(11):3402–3417. doi:10.1021/bi952626l
40. Nagata K, Kohda D, Hatanaka H, Ichikawa S, Matsuda S, Yamamoto T, Suzuki A, Inagaki F (1994) Solution structure of the epidermal growth factor-like domain of heregulin-alpha, a ligand for p180erbB-4. EMBO J 13(15):3517–3523