

Progress in Structural and Computational Studies of Epidermal Growth Factor Receptor*

Yonghui Zhang, Zhiyong Zhang[#]

Hefei National Laboratory for Physical Sciences at Microscale, School of Life Sciences,
University of Science and Technology of China, Hefei
Email: [#]zzyzhang@ustc.edu.cn

Received: Mar. 19th, 2013; revised: Apr. 3rd, 2013; accepted: Apr. 19th, 2013

Copyright © 2013 Yonghui Zhang, Zhiyong Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: Epidermal growth factor receptor (EGFR) plays a critical role in the network of signal transduction, and regulates cell proliferation, differentiation, and migration. EGFR is implicated in different kinds of cancers, which makes it an excellent target for anti-tumor therapy. This review considers structural and computational studies of EGFR, and their impact on our understanding of EGFR function. With continuing efforts of biologists, many high-resolution structures of different regions in EGFR, which include the extracellular region and the intracellular tyrosine kinase domain, have been solved. Starting from these structures, computer simulation techniques enable us to investigate large length-scale and long time-scale conformational changes during the EGFR activation. All the studies will be helpful to reveal the molecular mechanism in EGFR regulation, and provide valuable knowledge for developing anti-cancer drugs.

Keywords: Epidermal Growth Factor Receptor (EGFR); Structure and Function; Computer Modeling; Molecular Dynamics Simulations; Conformational Changes

表皮生长因子受体的结构与计算生物学研究进展*

张泳辉, 张志勇[#]

中国科学技术大学生命科学学院, 合肥微尺度物质科学国家实验室(筹), 合肥
Email: [#]zzyzhang@ustc.edu.cn

收稿日期: 2013年3月19日; 修回日期: 2013年4月3日; 录用日期: 2013年4月19日

摘要: 表皮生长因子受体(epidermal growth factor receptor: EGFR)在细胞信号转导中发挥了重要作用, 调节细胞的增殖、分化与迁移。EGFR 与多种癌症密切相关, 这使其成为一个理想的抗肿瘤药物靶点。本文对 EGFR 的结构生物学及计算生物学的研究进展做一综述。通过结构生物学家的不懈努力, 已经对 EGFR 的三维空间结构有了较为清晰的认识。计算机模拟是研究生物大分子动态过程的重要手段之一, 可以在 EGFR 静态结构的基础上, 模拟其在配体诱导的激活过程中发生的大尺度、长时程的动态构象变化。对 EGFR 的结构及其分子内部运动的研究, 有助于揭示其发挥生物学功能的分子机理, 为抗肿瘤药物的设计开发提供指导。

关键词: 表皮生长因子受体; 结构与功能; 计算机模拟; 分子动力学模拟; 构象变化

1. 引言

表皮生长因子受体 EGFR 属于受体酪氨酸激酶

*资助信息: 安徽省自然科学基金面上项目 1208085MC38; 高等学校博士学科点专项科研基金(新教师类)20113402120013。

(receptor tyrosine kinase: RTK), 该家族包括 4 个成员(称为 ErbB 或 HER 家族), 分别是 EGFR(ErbB1/HER1)、ErbB2(Neu/HER2)、ErbB3(HER3)和 ErbB4(HER4)。其中 EGFR 最早被克隆, 研究得最为充分。本

文首先介绍 EGFR 的生理功能及其在癌症研究中的重要意义。然后总结近年来对 EGFR 的结构解析取得的成果,在此基础上讨论高分辨率结构信息对解释 EGFR 的功能所起到的不可替代的作用。本文的另外一个重点是关于 EGFR 的计算机模拟研究,包括配体与受体的结合,以及配体诱导的 EGFR 二聚过程中发生的大尺度构象变化,等等。我们希望读者能了解基于结构的计算机模拟方法和技术在研究包括 EGFR 在内的生物大分子动态过程方面的应用潜力。最后,我们对 EGFR 相关的结构生物学和计算生物学研究做出展望。

2. 表皮生长因子受体的功能

EGFR 是位于细胞表面的受体,可以结合特异的配体,主要包括表皮生长因子(EGF)和转化生长因子 α (transforming growth factor α : TGF α),而被激活。当结合配体后,EGFR 单体发生二聚化,刺激其胞内酪氨酸激酶(tyrosine kinase: TK)活性,导致 C 端的一些酪氨酸残基(包括 Y992, Y1045, Y1068, Y1148 和 Y1173)自磷酸化。下游的信号蛋白通过自身 SH2 结构域与磷酸化的酪氨酸相互作用,从而激活下游信号通路(主要包括 MAPK、Akt 和 JNK 通路),引起级联放大效应。EGFR 调节细胞的增殖、分化与迁移,对人体皮肤的天然免疫反应非常重要。

癌症的发生通常与细胞非正常增殖有关,现已发现人体多种癌症的发展都与 EGFR 的突变或过量表达密切相关。EGFR 及其配体表达水平的升高已被认为是多种癌症的一般表现,并且有可能促进固体瘤的生长。Nicholson 等研究了 1985~2000 年间 PubMed 上的文献,分析 EGFR 的表达和癌症预报之间的关系^[1]。他们发现 EGFR 可在头部、颈部、卵巢、子宫颈、膀胱、食道癌的发生中作为有效的指示分子,在这些癌症中,上升的 EGFR 表达与复发率增加,以及存活率的减少有 70% 的关联。

因此,EGFR 是当前研究抗癌分子治疗的主要靶点之一,EGFR 信号通路也已成为研究得最为广泛的信号转导网络。全面透彻地理解 EGFR 的激活机制对发展新的癌症治疗方法具有重要的意义。有研究团队发现一些癌症治疗药物的敏感性和抗药性与 EGFR 的突变有关^[2],对这些突变体的研究将帮助我们寻找更

有效的抗癌药物。近年来,利用结构生物学的手段解析 EGFR 的高分辨率结构取得了重要进展,使我们可以更深入地了解 EGFR 激活的分子机理,研究 EGFR 突变对其结构的改变以及对抗癌药物分子的敏感性和抗药性的影响。

3. 表皮生长因子受体的结构生物学研究

人的全长 EGFR 有 1186 个氨基酸,由胞外区(extracellular region, 氨基酸序列 1-620),单一的跨膜域(transmembrane domain, 621-643),胞内的近膜区(juxtamembrane region, 644-685),酪氨酸激酶域(tyrosine kinase domain, 686-953)和一个长的柔性很大的 C 端调控区(C-terminal regulatory region, 954-1186)组成。胞外区由 4 个结构域(I-IV)组成,通常也称为 L1, CR1, L2 和 CR2 或者 L1, S1, L2 和 S2。结构域 I(1-165)和结构域 III(311-480)有 37% 的序列相似性,而结构域 II(166-310)和结构域 IV(481-620)富含半胱氨酸。

由于 EGFR 是一个分子量较大的跨膜蛋白,且局部结构柔性较大,所以目前还没有获得全长 EGFR 的晶体结构。现有 PDB 中已公布的 EGFR 相关结构都只包含全长的一部分,即 EGFR 的胞外区结构和胞内酪氨酸激酶区结构。这两个区域的结构都有非激活(图 1)和激活(图 2)两种状态。

3.1. EGFR 胞外区的结构

在 EGFR 胞外区(soluble extracellular region of EGFR: sEGFR)结构未获得解析之前,人们对配体(EGF 或 TGF α)诱导的 EGFR 二聚化方式并不清楚。绝大多数 RTKs,如血小板衍生生长因子(PDGF)/Kit 受体家族,配体自身是二聚化的,并使受体有效地交联成二聚复合物。对纤维生长因子(FGF)受体,附属分子使配体连结产生二价的复合物,然后使两个受体分子交联。2002 年 Cell 同期发表的两篇文章,分别解析了 sEGFR-EGF 和 sEGFR-TGF α 的二聚体晶体结构,彻底改变了人们对配体诱导的 EGFR 二聚化机制的认识^[3,4]。从这些结构可以看出(图 2),sEGFR 的结构域 I~III 成 C 型的排列,配体结合在结构域 I 和 III 之间,其结合位点和二聚的界面相隔一定距离因而对二聚无直接的贡献。二聚的界面是通过受体的结构域 II 维

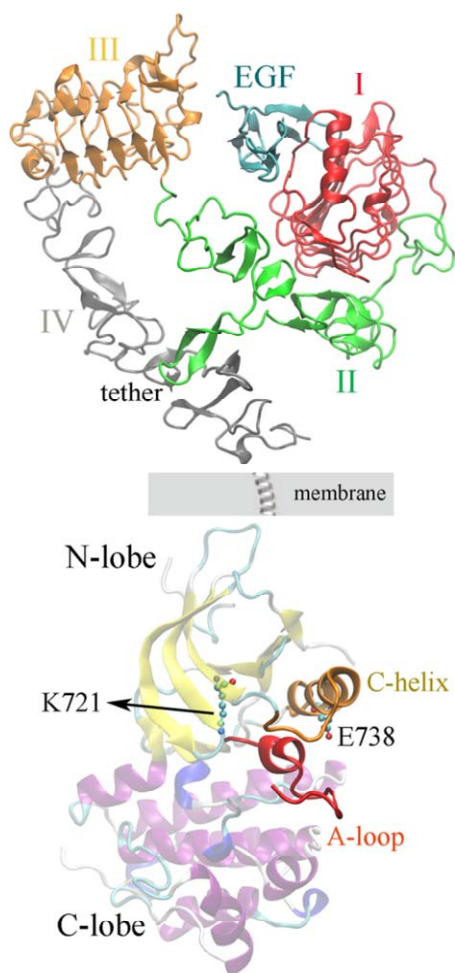


Figure 1. Inactive sEGFR and EGFR-TK
图 1. EGFR 胞外区和胞内酪氨酸激酶区非激活态构象

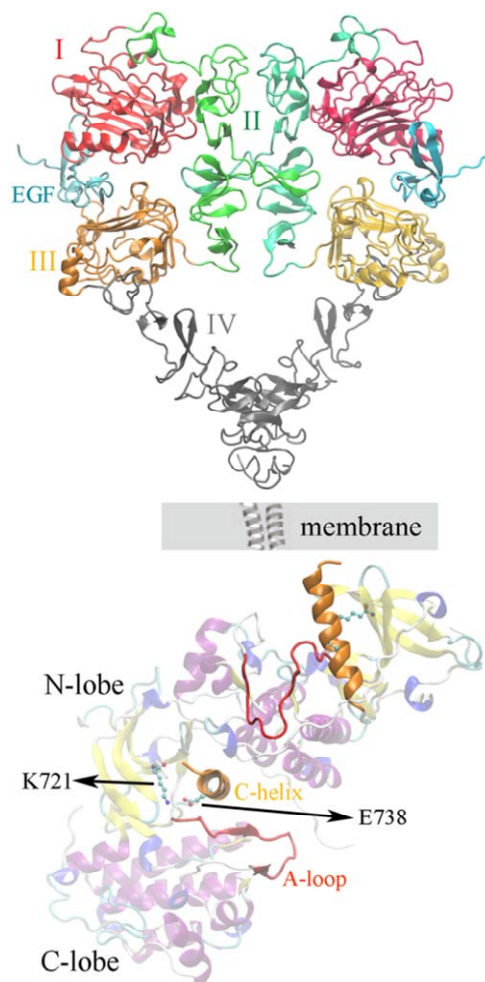


Figure 2. Active sEGFR dimer and EGFR-TK dimer
图 2. EGFR 胞外区和胞内酪氨酸激酶区激活态二聚体构象

系的。在结构域 II 里一个 β 发卡结构，也称聚合臂 (dimerization arm)，两个 EGFR 的聚合臂有很强的相互作用，其尖端甚至与对面 EGFR 分子的结构域 I 产生作用。

Ferguson 等解析了一个 sEGFR 非激活的晶体结构^[5]，他们发现 sEGFR 非激活态是一个自抑制的结构：其聚合臂因和结构域 IV 作用(称为 tethered interactions)而被包埋在分子内，而无法形成二聚体(图 1)。通过与 sEGFR 的激活构象比较，可以看出其在被激活前后发生了大尺度的构象变化。结构域 I 需要有大约 130° 的转动以及 20 \AA 的平动，sEGFR 才能从非激活态(tethered conformation)转变成激活态(extended conformation)。构象变化导致的主要后果是 tethered interactions 被破坏，以及二聚臂的暴露，从而使二聚变得可能。但是上述构象变化是如何发生的，以及

配体分子在其中所起的作用，现在仍然不清楚。

关于结构域 IV 在 EGFR 二聚中所起的作用，有人认为其对 sEGFR 二聚体的稳定有直接的贡献^[5,6]。但有实验发现，就算将整个结构域 IV 去除，对二聚体的结合强度影响甚微^[7]。2010 年，Lu 等解析了包括结构域 IV 在内的 sEGFR 二聚体的晶体结构^[8]，表明两个单体的结构域 IV 之间存在一个小的疏水接触面(图 2)，但是其对二聚的贡献尚待进一步研究。

PDB 中的 EGFR 相关结构绝大部分都是人源的，其胞外区二聚体表现出对称性。2010 年 Lemmon 研究组解析了果蝇的 2:2 sEGFR-EGF 晶体结构^[9]，发现两个 sEGFR 单体在结构上是非对称的，他们认为这是 EGFR 与配体结合负协同性的结构基础，即一个配体与 EGFR 结合后的构象变化会造成一个非对称的二聚受体，导致配体与剩下的一个结合位点亲和力减弱。

3.2. EGFR 胞内酪氨酸激酶区的结构

结合拉帕替尼(lapatinib)的 EGFR-TK 和结合 AMP-PNP 的 EGFR-TK 结构呈现出非活化状态^[10,11], 具有明显的分子内自抑制相互作用。两个结构非常相似而且与蛋白激酶 CDKs 和 Src 家族激酶也很像。在这两个结构中, activation loop(A-loop)中一个短的螺旋区紧靠起关键催化作用的 C-helix(图 1)。在非活化的 EGFR-TK、CDK 和 Src-家族激酶里, C-helix 通过与 A-loop 的相互作用而发生移动, 破坏二者之间相互作用的突变可以导致 EGFR-TK 活化。

活化的 EGFR-TK 晶体结构是以不对称的二聚体形式存在的^[11]。一个 EGFR-TK 分子的 C-lobe 紧靠着与它二聚的另一分子的 N-lobe, 导致在非活化构象中存在的 C-helix 和 A-loop 间的自抑制相互作用在活化的 EGFR-TK 二聚体中被破坏(图 2)。设计能打破不对称二聚体的突变可以阻断 EGF 对 EGFR 的激活。在 A-loop 中有一个保守的谷氨酸残基 E738, 它与赖氨酸 K721 形成盐键(图 2), 而在非活化的 EGFR-TK 中, 该盐键不存在(图 1)。

目前的观点认为 EGFR-TK 非对称二聚体激酶区活化的机理如下: 二体中一个 EGFR-TK 是 activator, 而另外一个 receiver。Receiver 分子的 N-lobe 与 activator 分子的 C-lobe 相互作用, 导致前者被激活, 从而磷酸化 activator 分子 C 端调控区的酪氨酸。然后两个 EGFR 互换角色, 最终使两个受体分子的胞内 C 端都完成磷酸化。

综合上述结构信息, 可以对 EGFR 活化的大致过程描绘如下: 在未结合配体时, EGFR 的胞外区和胞内酪氨酸激酶区都以自抑制的构象形式存在。配体的结合引起受体胞外区的二聚化, 使胞内区的距离拉近, 导致酪氨酸激酶区形成一个不对称的二聚体。该二聚体能够通过别构调节使激酶区活化, 从而使两个胞内 C 端区域的酪氨酸磷酸化^[12]。

虽然迄今已有很多 EGFR 的结构, 但一些重要区段还是很匮乏的。如可能有重要调控功能的胞内近膜区的前 30 个氨基酸, 以及含有多个酪氨酸磷酸化位点的 C 端 190 个氨基酸残基与受体激活的调控相关, 但有关的结构信息很少。

4. 表皮生长因子受体的计算生物学研究

虽然 EGFR 的结构信息已非常丰富, 但单从这些

静态结构我们还不能得知其行使功能时的动态变化。分子动力学(molecular dynamics: MD)模拟是研究在生理条件下生物大分子运动的有效手段, 它非常适合于揭示三级、四级结构的变化及其与相应生物功能的关系。虽然 EGFR 的分子量较大, 但随着近年来计算机软硬件的飞速发展, 对其进行 MD 模拟还是可以实现。因为 EGFR 的胞外区和胞内酪氨酸激酶区结构信息最为丰富, 目前有关 EGFR 的 MD 模拟均集中在这两个区域。

4.1. EGFR 胞外区的模拟

Kästner 等用 MD 模拟研究了 EGFR 胞外二聚体的空间取向、构象可塑性, 以及高聚的情况^[13]。与 EGFR 二聚体立于细胞膜上的传统的结构图不同, 他们发现二聚体仅需有限的重排, 即可几乎平躺在细胞膜上, 这和由荧光共振能量转移(FRET)和电镜实验得到的结果一致。与细胞膜的相互作用引起 EGFR 二聚体的构象的改变, 进而稳定结合界面的联系。他们还模拟了二聚体间的相互作用: 在细胞膜上, 两个二聚体能够通过一个“头对头”(head-to-head)界面形成一个四聚体, 其中涉及到胞外区的配体结合位点, 增强了配体的结合。在晶体结构中, “头对头”的界面也被发现, 但其在模拟中明显更加稳定。

与果蝇 sEGFR 二聚体的不对称结构相比, 人源 sEGFR 二聚体是对称结构, 其负协同性效应只在全长 EGFR 中才存在。Tynan 等基于定量 FRET 成像的方法, 结合蒙特卡洛和 MD 模拟, 以探测 EGFR 的构象^[14]。他们发现人的具有高配体结合能力的 sEGFR 构象是平躺在细胞膜上的。通过进一步研究这种构象与配体结合的负协同性间的联系, 发现该结构的不对称性与果蝇 sEGFR 的主要特点很相似。也就是说, 从无脊椎动物到人的 EGFR 负协同性的结构基础都是保守的, 但人源 EGFR 胞外区的不对称还需要与细胞膜的相互作用来维持。

我们对 sEGFR3-512(即结构域 IV 缺失)二聚体在溶液和结晶环境下进行了 MD 模拟^[15]。在水溶液中, 经过纳秒时间尺度的模拟后, sEGFR 的结构变得更加紧凑, 出现二体对称性的消失以及两个 sEGFR 单体间有更多的接触。而在晶体环境下的模拟, 由于晶格堆积作用使上述构象变化受到抑制。在模拟中缺失的结构域 IV 可能做为一个 spacer, 起到稳定二聚体的作

用。

4.2. EGFR 胞内酪氨酸激酶区的模拟

Papakyriakou 等对 EGFR-TK 的非活化和活化的结构进行了常规的 MD 模拟, 然后利用 targeted MD(TMD)模拟两种状态之间的构象转变^[16]。通过模拟他们发现 EGFR-TK 的二聚可以稳定那些与其活化态相关的结构要素, 并且预言了新的有 activation loop 参与的盐键。EGFR-TK 两种状态相互转化的 TMD 模拟证实了那些具有重要功能的保守结构区域的形成。对一个致癌突变体 L834R 的模拟揭示了其导致 EGFR 激活的可能结构基础: 1) 突变的 R834 残基与邻近酸性氨基酸残基形成盐键; 2) 导致非活化状态下的一个疏水簇不稳定。

Shan 等对野生型 EGFR-TK 及其突变体进行了微秒时间尺度的 MD 模拟, 发现野生型 EGFR-TK 的 N-lobe 二聚界面呈现无序状态, 但在形成二聚体后变得有序^[17]。作者认为一些在该区域与癌症相关的突变(特别是 L834R), 降低了该区域的无序性从而促进 EGFR 的二聚。他们还发现在 EGFR-TK 区的 Tyr845 的磷酸化可抑制其固有的无序性, 这一研究成果为 EGFR 自发信号转导机制提出了一种模型。

5. 展望

虽然当前 EGFR 的结构信息已相当丰富, 但其激活的很多细节仍不太清楚, 有许多尚未完成而又有重要意义的工作, 需要更多结构生物学和计算生物学的研究对 EGFR 的活化机制进行补充验证。例如, 我们需要解析全长 EGFR(包括单体和二聚体)尽可能完整的结构, 特别是胞内近膜区和 C 端调控区的结构信息。从分子模拟的角度, 可以进行以下工作: 1) 模拟配体诱导的 EGFR 胞外区从自抑制构象到伸展构象的转变路径; 2) 全长 EGFR 的多尺度模拟。我们可以根据目前已有的结构, 结合同源建模等方法构建全长 EGFR 的结构模型。然后利用多尺度计算机模拟方法(包括原子水平的 MD、粗粒化模拟等), 研究 EGFR 激活的动态变化过程, 在此基础上探讨其参与细胞信号转导的分子机理。

参考文献(References)

[1] R. I. Nicholson, J. M. Gee and M. E. Harper. EGFR and cancer

- prognosis. *European Journal of Cancer*, 2001, 37(Suppl 4): 9-15.
- [2] J. G. Paez, P. A. Janne, J. C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F. J. Kaye, N. Lindeman, T. J. Boggon, K. Naoki, H. Sasaki, Y. Fujii, M. J. Eck, W. R. Sellers, B. E. Johnson and M. Meyerson. EGFR mutations in lung cancer: Correlation with clinical response to gefitinib therapy. *Science*, 2004, 304(5676): 1497-1500.
- [3] H. Ogiso, R. Ishitani, O. Nureki, S. Fukai, M. Yamanaka, J. H. Kim, K. Saito, A. Sakamoto, M. Inoue, M. Shirouzu and S. Yokoyama. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell*, 2002, 110(6): 775-787.
- [4] T. P. J. Garrett, N. M. McKern, M. Z. Lou, T. C. Elleman, T. E. Adams, G. O. Lovrecz, H. J. Zhu, F. Walker, M. J. Frenkel, P. A. Hoynes, R. N. Jorissen, E. C. Nice, A. W. Burgess and C. W. Ward. Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor α . *Cell*, 2002, 110 (6): 763-773.
- [5] K. M. Ferguson, M. B. Berger, J. M. Mendrola, H. S. Cho, D. J. Leahy and M. A. Lemmon. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Molecular Cell*, 2003, 11(2): 507-517.
- [6] A. W. Burgess, H. S. Cho, C. Eigenbrot, K. M. Ferguson, T. P. J. Garrett, D. J. Leahy, M. A. Lemmon, M. X. Sliwkowski, C. W. Ward and S. Yokoyama. An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. *Molecular Cell*, 2003, 12(3): 541-552.
- [7] J. P. Dawson, M. B. Berger, C. C. Lin, J. Schlessinger, M. A. Lemmon and K. M. Ferguson. Epidermal growth factor receptor dimerization and activation require ligand-induced conformational changes in the dimer interface. *Molecular and Cellular Biology*, 2005, 25(17): 7734-7742.
- [8] C. Lu, L. Z. Mi, M. J. Grey, J. Zhu, E. Graef, S. Yokoyama and T. A. Springer. Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal growth factor receptor. *Molecular and Cellular Biology*, 2010, 30(22): 5432-5443.
- [9] D. Alvarado, D. E. Klein and M. A. Lemmon. Structural basis for negative cooperativity in growth factor binding to an EGF receptor. *Cell*, 2010, 142(4): 568-579.
- [10] E. R. Wood, A. T. Truesdale, O. B. McDonald, D. Yuan, A. Hassell, S. H. Dickerson, B. Ellis, C. Pennisi, E. Horne, K. Lackey, K. J. Alligood, D. W. Rusnak, T. M. Gilmer and L. Shewchuk. 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 Research*, 2004, 64(18): 6652-6659.
- [11] X. Zhang, J. Gureasko, K. Shen, P. A. Cole and J. Kuriyan. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell*, 2006, 125(6): 1137-1149.
- [12] K. M. Ferguson. Structure-based view of epidermal growth factor receptor regulation. *Annual Review of Biophysics*, 2008, 37: 353-373.
- [13] J. Kastner, H. H. Loeffler, S. K. Roberts, M. L. Martin-Fernandez and M. D. Winn. Ectodomain orientation, conformational plasticity and oligomerization of ErbB1 receptors investigated by molecular dynamics. *Journal of Structural Biology*, 2009, 167(2): 117-128.
- [14] C. J. Tynan, S. K. Roberts, D. J. Rolfe, D. T. Clarke, H. H. Loeffler, J. Kastner, M. D. Winn, P. J. Parker and M. L. Martin-Fernandez. Human epidermal growth factor receptor (EGFR) aligned on the plasma membrane adopts key features of Drosophila EGFR asymmetry. *Molecular and Cellular Biology*, 2011, 31(11): 2241-2252.
- [15] Z. Zhang, W. Wriggers. Polymorphism of the epidermal growth factor receptor extracellular ligand binding domain: The dimer interface depends on domain stabilization. *Biochemistry*, 2011, 50(12): 2144-2156.
- [16] A. Papakyriakou, D. Vourloumis, F. Tzortzatos-Stathopoulou and M. Karpusas. Conformational dynamics of the EGFR kinase

- domain reveals structural features involved in activation. *Proteins*, 2009, 76(2): 375-386.
- [17] Y. Shan, M. P. Eastwood, X. Zhang, E. T. Kim, A. Arkhipov, R. O. Dror, J. Jumper, J. Kuriyan and D. E. Shaw. Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. *Cell*, 2012, 149(4): 860-870.