Computational modeling of an EGFR single-mutation resistance to cetuximab in colorectal cancer treatment

Ignasi Buch, Noelia Ferruz, Gianni De Fabritiis*

Computational Biophysics Laboratory (GRIB-IMIM), Universitat Pompeu Fabra, Barcelona Biomedical Research Park (PRBB), C/Doctor Aiguader 88, 08003 Barcelona, Spain.

Supporting Information Placeholder

ABSTRACT: Extracellular S468R mutation of the epidermal growth factor receptor (EGFR) was recently identified as the cause of resistance to cetuximab, a widely used drug in colorectal cancer treatment. Here, we have determined the binding free energies of cetuximab's Fab V_H-V_L domains and endogenous EGF ligand to wild type and S468R EGFR by high-throughput molecular dynamics. The results show that the mutation only marginally affect the affinity of the drug and suggest that increased competition between the drug and the endogenous ligands could be the reason to S468R EGFR resistance to cetuximab. Also, we have performed a comparative structure-based study of the binding of endogenous EGFR ligands with cetuximab and necitumumab antibodies. This work provides a interesting possible mechanism of resistance for this single mutation on such a large interaction surface which can be further tested experimentally.

Colorectal cancer is the third-leading cause of cancer-related deaths worldwide, with over 600,000 deaths occurring each year.[1] Recently, a role has been established for the epidermal growth factor receptor (EGFR, Her1) signal transduction pathway in the development of a subset of epithelial tumors.[2] EGFR is one of the four members of the Her (ErbB) family of receptors (Her1-4), directly involved in multiple cellular proliferation processes, including growth, differentiation, migration, and apoptosis. More concretely, EGFR overexpression has been shown to predict tumor progression[3] in colorectal cancer and is overexpressed in 25-77% of these tumors.[4]

In recent years, great efforts have been put on understanding EGFR activation and interactions[5-7] and different EGFR-targeting agents have been developed. The two agents that have demonstrated the best responses are two monoclonal antibodies directed against EGFR: cetuximab and panitumumab[8] (known as anti-EGFR therapy or EGFR inhibitors) which compete against endogenous EGFR ligands such as EGF, for both binding site and block receptor dimerization[9] (see Figure 1). These antibodies have presented high responses when administered with chemotherapy.

Cetuximab is a chimeric (mouse/human) IgG1 anti-EGFR monoclonal antibody that has demonstrated antitumor activity in patients with colorectal cancer.[10] However, cetuximab's murine structural component is a potential source of toxicity and immunogenicity[11] and a considerable amount of research aimed at eliminating this toxicity has been performed. As a result, a new agent was developed: panitumumab, a fully human IgG2 monoclonal antibody that is highly selective for EGFR.[8] Both cetuximab and panitumumab are considered equivalent in the treatment of colorectal cancer and therefore it was assumed that both share the same epitope.[12,13]

However, a missense mutation has been identified in the extracellular domain III of EGFR, S492R (In this work, we will refer to the mutation as S468R, according to residue numbering in crystal structure Li et al.[9]: antigen binding fragment from cetuximab (FabC255) in complex with the soluble extracellular region of EGFR). Whereas the mutation has been identified as the cause for acquired resistance to clinical treatment of colorectal cancer with cetuximab, this has not been the case for panitumumab.[13]

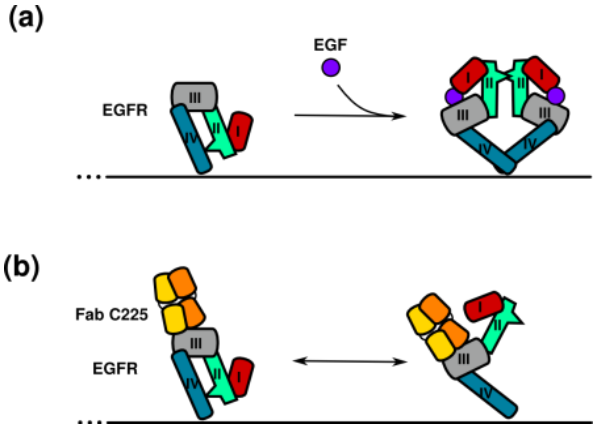

Figure 1. Schematic binding of endogenous ligand EGF and FabC255 antibody. (a) After EGF binding to I and III domains, EGFR receptor adopts a conformation compatible with dimerization. (b) FabC255 binds by interacting only with domain III and impeding both dimerization of EGFR and further inactivating the receptor.

Several mutations in domain III in EGFR have been previously reported in the literature that help to understand the role of epitopic residues in the binding of cetuximab, but the study by Montagut et al.[8] is the first example of a missense mutation of the target of an antibody being the direct cause of resistance. However, the molecular basis of this resistance remained unclear and the S468R mutated receptor was con-

cluded to be selectively defective in binding cetuximab.[13] Understanding the molecular basis of such interaction mechanism could clearly lead to a better understanding of the chemical role of the mutation and the development of more effective targeted therapies or new therapeutic combinations.[12]

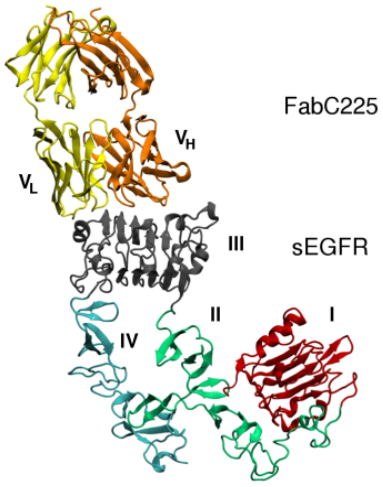

Figure 2. Complex structure for the bound FabC225 (antigen binding fragment of cetuximab) and single chain EGFR as crystalized by Li et al.[9] Computational modeling was performed on the binding domains of both partners, V_L and V_H for FabC225 and domain III for EGFR.

For this reason, in order to determine the cause of an acquired resistance at molecular level, we have modeled both cetuximab-EGFR and EGF-EGFR complexes for mutant and wild type EGFR and subsequently performed computational binding free energy calculations.[14,15] Calculating binding free energies using molecular dynamics simulations (MD) is a widely explored topic in the field of computational biophysics and several methodologies have been successfully developed.[11,16-18] Here, we apply a previously described umbrella sampling protocol for all-atom binding free energy calculations of large and semi-flexible protein-protein complexes.[15] Namely, we computed the binding free energies of both antigen and endogenous ligand to mutant and *wt* EGFR, as explained in Supplementary Information. Available crystal structures of FabC255-EGFR (antigen binding fragment of cetuximab) and EGF-EGFR crystals were used for MD simulations,[9,19] where only the interacting part of EGFR (domain III) was included in both models and the interacting part of V_H-V_L domains in FabC255 in the case of this antibody (Figure 2). Consequently, four complexes were modeled: wild type antigen and ligand receptor complex and their respective S468R mutants (from here on systems will be referred as the *wt* and S468R FabC255 V_H-V_L-EGFR and EGF-EGFR systems, respectively). S468R mutants of EGFR were generated on the basis of the *wt* EGFR crystallographic model. Systems were then parametrized, equilibrated and simulated using ACEMD[20] on GPUGRID.net[21] (See supporting information for further details).

Analyzing 25 μs of aggregated simulation time using the protocol described in Buch et *al.*[15] we obtained the binding affinities for the four complexes. Results reported in Figure 3a are the mean and standard deviation of the 5 replicas per system where each replica value is obtained from the latest quarter of sampled time in FabC255 V_H-V_L, and from the last 10-20 ns sampled in EGF (see supplementary information for further data). The calculated free energy values only considered interactions with domain III of EGFR, and are directly comparable to the experimental values reported in Figure 3a.

EGF were found to bind to EGFR domain III with a free energy of -6.8 ± 0.5 kcal/mol, compared to an experimental of -7.7kcal/mol.[9] This measure is pretty accurate with less than 1 kcal/mol difference from the experimental value, being this level of approximation in line with a previous work for a tetrapeptide ligand using the same protocol.[15,22] Considering the size of the system, the accuracy of these calculations for the *wt* complexes is remarkably high. The same calculations for the simplified FabC255 V_H-V_L system gave -9.8 ± 0.3 kcal/mol with a difference of 2.1 kcal/mol compared to the experimental value of -11.9 kcal/mol to sEGFR domain III.[9] This larger difference in binding affinity may be explainable by the lack of the V_H-V_L domains far away from EGFR not considered in the computational model, yet as the mutation is totally embedded in the interface between domain III and the antibody, its effect is well accounted.

(a)

	wt (exp)	wt (comp)	S468R	$\Delta\Delta G^0_{mut-wt}$
EGF	-7.7 ± 0.1	-6.8 ± 0.5	-7.9 ± 0.6	-1.1 ± 0.8
FabC255 V_H-V_L	-11.9 ± 0.1	-9.8 ± 0.3	-8.8 ± 0.4	-1.0 ± 0.5

(b)

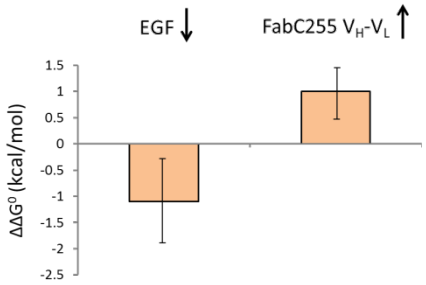

Figure 3. **(a)** Table showing binding free energies calculations (ΔG^0) from experimental measurements (exp.) and computational calculations (comp.) for the EGF and FabC255 systems to the *wt* and S468R structures of EGFR domain III. All units are in kcal/mol. **(b)** Plot showing the differences in binding affinity between mutant and wild type systems in EGF and FabC255. EGFR domain III S468R mutation increases the affinity for endogenous ligand EGF and decreases the affinity for FacC225 V_H-V_L, according to the formula: $\Delta\Delta G = \Delta G(mutant)-\Delta G(wt)$. Error bars are depicted in the plot.

Analysis of the S468R mutant complexes provided interesting results. Affinity of FabC255 V_H-V_L does not change significantly and suggest that there must be other mechanism of resistance different that total loss of affinity. Actually, calculations predicted a net strengthen of EGF ligand of 1.1 ± 0.5 kcal/mol and a mild weakening of FabC255 V_H-V_L of 1 ± 0.8 kcal/mol. (Figure 3b).

In quantitative terms, these results show a change on the bound populations between wild type and mutant receptor. In the former case, only a 0.4% of the receptors would be bound to EGF according to differences in binding energies

($\Delta\Delta G_{Fabc255\text{-}EGF}$ = -3.3kcal/mol). In the case of the mutant, a 22% of the receptors would be interacting with EGF, given concentrations of endogenous ligand and antibody are both found in the same nanomolar range *in vivo*.[23-25]

Furthermore, EGF also interacts with domain I and is known to bind with an affinity almost 2kcal/mol greater to full sEGFR [9,10] (-9.4kcal/mol). Same increment in binding affinity should be expected for the full S468R receptor, but FabC255 solely interact with domain III of sEGFR and differences in affinities between drug and ligand are expected to be even more remarkable.

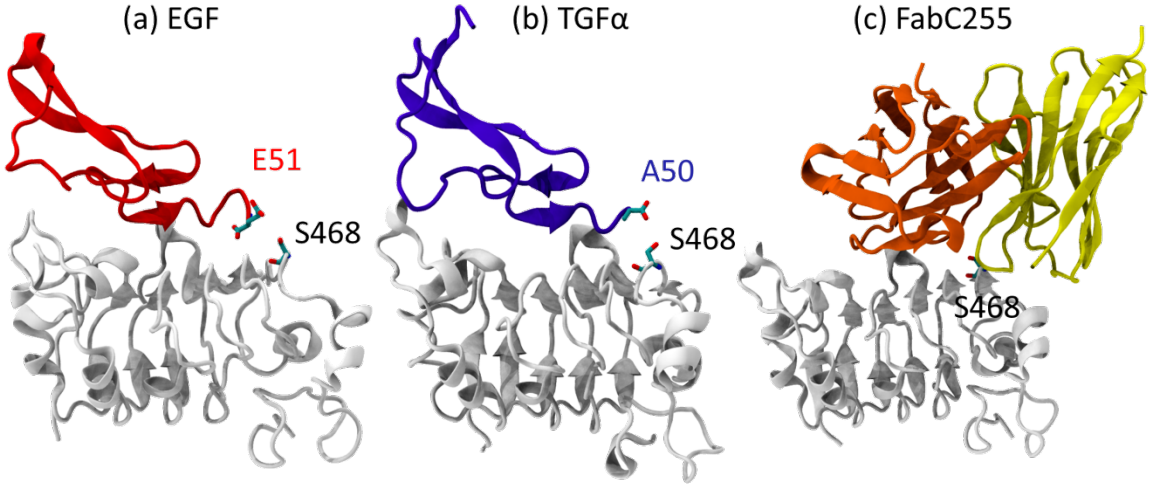

Figure 4. Spatial relationship between the mutation site S468 in EGFR and the bound structures of FabC225 and other receptor-activating ligands. **(a)** In the case of EGF, the mutation may increase the affinity of the ligand. A new salt bridge interaction may exist upon mutation of serine to arginine with E51 in EGF. **(b)** TGFα has A50 in close proximity to mutant site. Binding affinity might be slightly increased after mutation S468R by interaction between the C-terminal of TGFα and arginine's side chain. **(c)** S468R mutation may have an impact on the binding of FabC225. A complex number of surrounding interactions may be affected with the inclusion of a large and charged amino acid as it is shown.

In order to understand in greater detail the chemical mechanism for this inversion in affinity, we have performed a computational molecular structure-based assessment to study the impact of the S468R EGFR mutation.

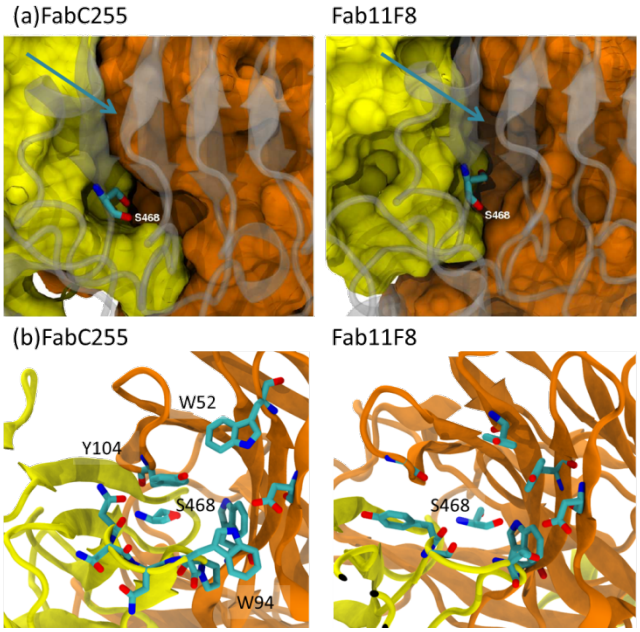

Figure 5. Assessing the effect of mutation S468R in FabC255 (cetuximab antibody fragment binding) and Fab11F8 (necitumumab antibody fragment binding) **(a)** Compared to FabC225, Fab11F8 displays a cavity that may be able to accommodate a bulky residue such as arginine. **(b)** A selection of residues between 5Å and 7Å of S468 in the crystal structures show the residues responsible for the different antibody surfaces. Principal differences between FabC225 and Fab11F8 are residues Y104, W52 and W94 in FabC225. Y104 in FabC225 practically obstructs an otherwise accessible cavity for arginine. Tryptophan side chains parallel to the binding interface contribute to a strong hydrophobic interaction.

Endogenous ligands of EGFR were the ligands considered in the study (EGF and transforming growth factor alpha, TGFα), whereas FabC255 domain of cetuximab antibody and an alternative anti-EGFR human monoclonal antibody named necitumumab were the considered antibodies.

The crystallographic structure of FabC225 in complex with the soluble extracellular sEGFR shown in Figure 2 reveals a single interaction interface between the drug and the target. The interface is similar to that of the endogenous EGFR ligands which makes FabC255 a competitive inhibitor.[9] Modification of these interfaces has the potential to affect complex formation and, as shown in Figure 4, the new S468R mutation may have a different impact for the binding of the drugs or the ligands.

The site of the S468R mutation lies right in the middle of the surface recognized by the FabC255 domain as shown in Figure 4c. In the case of EGF and TGFα, the mutation is placed very close to the C-terminal and whereas an additional salt bridge may become possible between E51 and S468 for the case of EGF, the mutation is not expected to have any effect on the binding affinity of TGFα (Figure 4a-b). For HRGα (heregulin), a non-native ligand of EGFR but activator of closer proteins in the EGFR family,[26] the mutation might presumably seat right underneath the interface of binding, and it would be expected to favourably interact with E57 and E61. Accordingly, HRGα-mediated resistance to cetuximab was recently reported.[27]

There is no currently available crystal structure for panitumumab, but the crystal structure of an alternative anti-EGFR antibody, necitumumab, (Fab11F8) was published in 2008.[20] Therefore, we have visually assessed the impact of the S468R mutation on the FabC255-EGFR interface compared to that of EGFR-Fab11F8 (Figure 5). Despite having very similar epitopes, the complexity of the interaction interfaces is such that a simple mutation may have consequences for the affinity of the complexes, as already inferred for panitumumab.[13] The missense S468R mutation involves a change from a rather small, polar and uncharged side chain in serine to a large and electrically charged side chain in arginine. These steric and electrostatic changes combined, may have deleterious effects in maintaining tight hydrophobic interactions and shape-complementary in protein-protein interfaces. As shown in Figure 5b, the principal differences between the two antibodies are the presence of residues Y104, W52 and W94 in FabC255 bound near the S468 in EGFR. Residue Y104 in particular, appears to be obstructing an otherwise accessible cavity for arginine. Overall, the cost in a major side chain rearrangement in cetuximab residues near S468 may be incompatible with the conservation of a tightly bound antibody. Moreover, the addition of a charged amino acid and a slight separation of the complex may facilitate the entrance of water molecules and complicate the tight complex formation characteristic of antigen-antibody interfaces. In the same way, Figure 4a represents both antigen surfaces and shows a bigger cavity which would be compatible with the accommodation of a bulkier residue as Arginine, displaying therefore a less negative response to the mutation.

None of the reported mutations in vitro[9,28] can individually match the deleterious effect that acquired mutation S468R promotes in cetuximab binding. Although highly significant for a single residue mutation, it alone is not enough to totally disrupt the interaction of the drug on such a large interaction surface. Our results show that the increased drug resistance may result from the combination in EGF binding affinity and the putative competition by other EGF-like ligands like HRGα that impedes receptor inhibition in vivo. Most of the mutations that have been experimentally explored in EGF binding to EGFR domain III actually caused a decrease in affinity.[28] Only the combined mutation Q408A/H409A has a significant increase of 2.7-fold in EGF binding affinity.[28] In this work, we show how a single missense mutation can cause both a mild decrease in drug binding and an increase in endogenous ligand binding, which may be the reason of the failure of the therapeutic strategy versus S468R mutant EGFR.

The affinity between cetuximab and the domain III is not drastically perturbed by the mutation, even considering potential approximations of the force field. Therefore, these results show that the competition mechanism should be the reason of the resistance, perfectly explaining the experimental findings. Indeed, experimental flow cytometry binding studies suggested that only panitumumab was able to bind to mutant EGFR expressing cells,[13] but it should be noted that these experiments included whole cells and endogenous and other ligands could be successfully competing against an unperceived lower affinity cetuximab–EGFR complex.

Finally, this work is an example of how in the near future and in the context of personalized medicine, binding free energy calculations could be successfully used to predict the efficacy of existing drugs to unknown target variants, and is expected to increasingly become routine in the years to come.

ASSOCIATED CONTENT

Supporting Information

Details of the equilibration, simulation protocols and binding energy calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

gianni.defabritiis@upf.edu

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENTS

IB acknowledges support from the Fundació La Marató de TV3. NF acknowledges support from Generalitat de Catalunya. GDF acknowledges support by the Spanish Ministry of Science and Innovation (Ref. BIO2011-27450). We also thank all the volunteers of GPUGRID who donate GPU computing time to the project. We finally thank C. Montagut for useful discussions on the role of the mutation.

REFERENCES

1. WHO (2012) Who cancer fact sheet. *http://tinyurl.com/6fg6ofz (last visited Nov. 15, 2013)*

2. McKay, J. A. *et al. Eur. J. Cancer Oxf. Engl. 1990* **38,** 2258–2264 (2002).

3. Porebska, I., Harlozińska, A. & Bojarowski, T. *Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med.* **21,** 105–115 (2000).

4. Hoy, S. M. & Wagstaff, A. J. *Drugs* **66,** 2005–2014; discussion 2015–2016 (2006).

5. Endres, N. F. *et al. Cell* **152,** 543–556 (2013).

6. Arkhipov, A. *et al. Cell* **152,** 557–569 (2013).

7. Shan, Y., Arkhipov, A., Kim, E. T., Pan, A. C. & Shaw, D. E. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 7270–7275 (2013).

8. Yang, X. D., Jia, X. C., Corvalan, J. R., Wang, P. & Davis, C. G. *Crit. Rev. Oncol. Hematol.* **38,** 17–23 (2001).

9. Li, S. *et al. Cancer Cell* **7,** 301–311 (2005).

10. Prewett, M. *et al. J. Immunother. Emphas. Tumor Immunol. Off. J. Soc. Biol. Ther.* **19,** 419–427 (1996).

11. Klastersky, J. *Curr. Opin. Oncol.* **18,** 316–320 (2006).

12. Bardelli, A. & Jänne, P. A. *Nat. Med.* **18,** 199–200 (2012).

13. Montagut, C. *et al. Nat. Med.* **18,** 221–223 (2012).

14. Doudou, S., Burton, N. A. & Henchman, R. H. Standard Free *J. Chem. Theory Comput.* **5,** 909–918 (2009).

15. Buch, I., Sadiq, S. K. & De Fabritiis, G. *J. Chem. Theory Comput.* **7,** 1765–1772 (2011).

16. Woo, H.-J. & Roux, B. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 6825–6830 (2005).

17. Gervasio, F. L., Laio, A. & Parrinello, M. *J. Am. Chem. Soc.* **127,** 2600–2607 (2005).

18. Jorgensen, W. L. *Science* **303,** 1813 –1818 (2004).

19. Ogiso, H. *et al. Cell* **110,** 775–787 (2002).

20. Harvey, M. J., Giupponi, G. & Fabritiis, G. D. *J. Chem. Theory Comput.* **5,** 1632–1639 (2009).

21. Buch, I., Harvey, M. J., Giorgino, T., Anderson, D. P. & De Fabritiis, G. *J. Chem. Inf. Model.* **50,** 397–403 (2010).

22. Giorgino, T., Buch, I. & De Fabritiis, G. *J Chem Theory Comput* **8,** 1171–1175 (2012).

23. Joh, T. *et al. Clin. Chim. Acta Int. J. Clin. Chem.* **158,** 81–90 (1986).

24. Erbitux. *http://www.accessdata.fda.gov/drugsatfda_docs//label/2004/125084lbl.pdf (last visited Nov. 15, 2013)*

25. Jean, G. W. & Shah, S. R. *Pharmacotherapy* **28,** 742–754 (2008).

26. Jacobsen, N. E. *et al. Biochemistry (Mosc.)* **35,** 3402–3417 (1996).

27. Yonesaka, K. *et al. Sci. Transl. Med.* **3,** 99ra86 (2011).

28. Li, S., Kussie, P. & Ferguson, K. M. *Struct. Lond. Engl. 1993* **16,** 216–227 (2008).

TOC GRAPHIC