Available online at www.sciencedirect.com
ScienceDirect
Computational Biology and Chemistry
Computational Biology and Chemistry 31 (2007) 134–137
www.elsevier.com/locate/compbiolchem

Brief communication

Computational simulation of interactions between SARS coronavirus spike mutants and host species-specific receptors

Yuan Zhang [a,*], Nan Zheng [b], Peng Nan [a], Ying Cao [c,d], Masami Hasegawa [c,d], Yang Zhong [a,e]

[a] School of Life Sciences, Fudan University, Shanghai 200433, China
[b] Division of Viral Immunology, Center for AIDS Research, Kumamoto University, 2-2-1 Honjo, Kumamoto 860-0811, Japan
[c] The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan
[d] Department of Biosystems Science, The Graduate University for Advanced Studies, Hayama, Kanagawa 240-0193, Japan
[e] Shanghai Center for Bioinformation Technology, Shanghai 200235, China

Received 23 November 2006; accepted 13 February 2007

Abstract

As a critical adaptive mechanism, amino acid replacements on the severe acute respiratory syndrome coronavirus (SARS-CoV) spike protein could alter the receptor-binding specificity of this envelope glycoprotein and in turn lead to the emergence or reemergence of this viral zoonosis. Based on the X-ray structures of SARS-CoV spike receptor-binding domain (RBD) in complex with its functional receptor (angiotensin-converting enzyme 2, ACE2), we perform computational simulations of interactions between three representative RBD mutants and four host species-specific receptors. The comparisons between computational predictions and experimental evidences validate our structural bioinformatics approaches. And the predictions further indicate that some viral prototypes might utilize the rat ACE2 while rats might serve as a vector or reservoir of SARS-CoV.

Keywords: SARS coronavirus; Spike protein; ACE2; Receptor-binding specificity; Structural bioinformatics

1. Introduction

Will the severe acute respiratory syndrome (SARS) reemerge in the coming years? This is a huge challenge of great concerns facing the whole scientific community especially virologists. As an envelope glycoprotein on the virion surface, the spike protein of SARS coronavirus (SARS-CoV) plays a crucial role in the virus entry (Hofmann and Pohlmann, 2004). This structural protein contains two functional regions: the external S1 region responsible for the initial attachment to cellular receptor and the internal S2 region contributing to the subsequent fusion between viral envelope and cellular membrane. A membrane-associated zinc metallopeptidase, angiotensin-converting enzyme 2 (ACE2), was identified as the functional receptor for SARS-CoV (Li et al., 2003). And a 193-amino acid (residues 318–510) independently folded receptor-binding domain (RBD), harbored within the S1 region, could efficiently bind human ACE2 (Wong et al., 2004). Moreover, the observation that the murine ACE2, compared to its human homologue, highly limited the efficient replication of SARS-CoV human isolate TOR2 in mouse cells strongly suggested the presence of species-specific viral entry barriers and the absence of post-entry barriers. Furthermore, two naturally occurred amino acid substitutions (K479N and S487T) were experimentally confirmed to be the critical determinants for the transmission of this zoonosis from civet cat to human (Li et al., 2005b). The recently solved crystal structures of the spike RBD in complex with its receptor, human ACE2, also demonstrated the significance of those two residue replacements in the adaptation of viral spike to human receptor (F. Li et al., 2005; Li et al., 2006). Clearly, the ACE2–RBD complex structure would undoubtedly facilitate the researches on the receptor recognition of SARS-CoV and the adaptive mechanism of virus crossing species barriers. And the recently developed structural bioinformatics program (Liu and Kuhlman, 2006) offer us a powerful tool to elucidate structural and energetic factors involved in SARS emergence or reemergence. In this study, we conducted structure-based computational simulations of interactions between three repre-

* Corresponding author. Tel.: +86 21 55664436; fax: +86 21 65642468.
E-mail address: yuanzhang5@fudan.edu.cn (Y. Zhang).

1476-9271/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiolchem.2007.02.006

sentative SARS-CoV RBD mutants (TOR2, GD and SZ3) and four host species-specific receptors (human, civet cat, mouse and rat) for assessment of their performance to predict viral crossing-barrier events.

2. Materials and methods

We retrieved the coordinates data of SARS-CoV spike RBD complexed with human ACE2 peptidase domain (F. Li et al., 2005) (PDB code: 2AJF) from the protein data bank (PDB) (Berman et al., 2000). There are two complexes in the asymmetric unit of crystal (F. Li et al., 2005). The structural differences between the two complexes are mainly due to the conformational changes of cellular receptors rather than those of viral RBDs. As a membrane-associated vasopeptidase, ACE2 had a shell-like catalytic extracellular domain (Towler et al., 2004), and its active site is deeply buried in a long cleft whose two sides are just the two lobes of the shell. After the engagement of substrate or inhibitor, the two lobes will exhibit a large hinge-bending movement toward each other to put relevant residues into right positions for catalysis. In the coordinates file, one ACE2–RBD complex (the complex AE with the chain A for ACE2 and the chain E for RBD) contains a fully open receptor while the other (the complex BF with the chain B for ACE2 and the chain F for RBD) has a slightly closed counterpart. Thus, as revealed in previous studies (F. Li et al., 2005; Towler et al., 2004), the former enzyme was in a free form whereas the latter in an inhibitor-bound state. However, both biochemical and structural researches (F. Li et al., 2005; Li et al., 2005b; Wong et al., 2004) indicated that the human receptor-binding activity of RBD is independent of this open-to-closed conformational transition. Hence, the two complex structures are respectively subject to structural bioinformatics analyses for further comparison. The in silicon analyses included three steps. Firstly, the contact map analysis (CMA) (Sobolev et al., 2005) was employed to define the contacting residues across the complex interfaces. Here, we focused on the two RBD residues (ASN479 and THR487) because their importance in the adaptation of SARS-CoV spike protein to human ACE2 has been biochemically and structurally clarified (F. Li et al., 2005; Li et al., 2005b). Additionally, those spike or ACE2 residues forming side chain contacts with the two focused positions are also under study due to their potential roles involved in the geometric and chemical complementarity of binding surface. After the definition of contacting residues, the RosettaDesign web server (Liu and Kuhlman, 2006) is consulted to redesign the interactions between RBD and receptor. In this step, the RBD positions 479 and 487 plus their contacting sites on the two binding partners are virtually mutated to mimic the contacts between the RBDs of three representative viral isolates and the ACE2 of four host species (F. Li et al., 2005; Li et al., 2005b). The three viral strains includes TOR2 (GenBank Accession No.: AY274119), GD (GenBank Accession No.: AY525636) and SZ3 (GenBank Accession No.: AY304486) which are isolated from the human patient during the serious pandemic in late 2003 (Marra et al., 2003; Rota et al., 2003), from the human patient of the mild infection case in the reemergence of 2004 (Consortium, 2004), and from the civet cat in the exotic market in early 2003 (Guan et al., 2003), respectively. Four host species comprised human (GenBank Accession No.: Q9BYF1), civet cat (GenBank Accession No.: Q56NL1), mouse (GenBank Accession No.: Q8R0I0) and rat (GenBank Accession No.: Q5EGZ1). Both the viral isolates and the host species were selected corresponding to previous virological studies (F. Li et al., 2005; Li et al., 2004, 2005b). As shown in Tables 1 and 2, the RBD positions 479 and 487 plus their contacting ACE2 sites of species-specific residue are redesigned according to the sequence content variations of viral isolates and host receptors identified as crucial determinants for receptor recognition (F. Li et al., 2005; Li et al., 2006, 2005b). Those residues, within RBD or ACE2, forming side chain contacts with the redesigned residues, are also repacked to accommodate the structural variations. Except the two categories of residues to be redesigned or repacked, all the others remained untouched in the complex structure. These computational simulations could serve as a reliable indicator for the validation of our structural bioinformatics approaches in this case, depending on its reproduction of receptor-binding pattern generated in the experimental investigations. Finally, the redesigned ACE2–RBD complex structure models are submitted to the FOLDX web server (Schymkowitz et al., 2005), an online empirical force field, to assess the effect of mutations on the binding free energy of complex and to measure the differences in the association energy between the wild type and the mutants. In principle, the lower the binding free energy, the higher the binding affinity.

Table 1
Redesigned residues of three representative viral isolates spike receptor-binding domain

	479	487
TOR2	N	T
GD	N	S
SZ3	K	S

Table 2
Redesigned residues of four host species ACE2 receptor

	31	34	353
Civet cat	T	Y	K
Human	K	H	K
Mouse	N	Q	H
Rat	K	Q	H

3. Results and discussion

Structural bioinformatics analyses of the ACE2–RBD complexes were performed using a three-step procedure as described in Section 2. The calculated binding free energies of the wild type and virtual mutant complex structures are shown in Table 3. Here, the agreement or disagreement between the computational predictions and the experimental evidences is paid special attention for the validation of our structural bioinformatics

Table 3
Calculated binding free energies of the wild type and virtual mutant ACE2–RBD structures based on the sister complexes AE and BF

	GD	SZ3	TOR2
Complex AE			
Civet cat	−13.2	−11.8	−14.5
Human	−14.1	−10.9	**−15.8**
Mouse	−12.0	−12.3	−12.3
Rat	−12.0	−14.4	−12.3
Complex BF			
Civet cat	−12.2	−14.5	−17.7
Human	−10.7	−12.3	**−14.0**
Mouse	−12.1	−14.0	−12.4
Rat	−12.1	−14.0	−12.4

The unit for binding free energy is kcal/mol. The viral isolates include GD, SZ3 and TOR2. The host species comprise civet cat, human, mouse and rat. The energy values of wild type complexes are shown in boldface.

approaches and the discrimination of predictability between the sister complexes AE and BF. For instance, both pseudovirus infections and binding assays reveals that all the three RBDs mentioned above could attach to the civet cat receptor whereas only the TOR2 RBD is able to efficiently utilize the human ACE2 (Li et al., 2005b). In addition, the TOR2 RBD is shown to lack the rodent receptor-binding activity (Li et al., 2004, 2005b). As for the models derived from the AE complex, the civet cat ACE2 attaches to the RBDs of TOR2, GD and SZ3 with the binding free energies in an ascending order (the binding affinities in a descending order), being congruent with the observations in binding assays (Li et al., 2005b). And a similar rank of RBD association energies is also found for the human receptor, completely consistent with the data of pseudovirus infections and binding assays (Li et al., 2005b). In particular, the wild type complex AE (TOR2–human), representing the interactions between human receptor and the human-isolated TOR2 RBD obtained during the serious emergence of SARS in 2003, holds the lowest energy (−15.8 kcal/mol) among the 12 AE-based complexes. While the highest binding energy score (−10.9 kcal/mol) is rewarded to the mutant complex (SZ3–human) mimicking the contacts of civet cat-isolated SZ3 RBD with human ACE2. In the middle of the two extremes lies in the energy value of the mutant complex (GD–human) modeling associations between human ACE2 and the human-isolated GD RBD in the reemerging case of SARS with mild symptoms. The significant differences in the human receptor-binding energy or affinity among the three viral RBDs, especially that between the two extremes, offer us an excellent interpretation for the human receptor-binding specificity differentiation of SARS-CoV spike protein given the high correlation with biochemical evidences (Li et al., 2004, 2005b). In addition, the equally higher binding energies (−12.3 kcal/mol) of rodent ACE2 with TOR2 RBD successfully account for the viral entry barriers posed by the rodent receptors (Li et al., 2004). Collectively, the virtual substitutions strongly suggest the validity of our structural bioinformatics protocols and the predictability of AE complex. Notably, this directly leads to the following speculation that rats might be a vector or reservoir for some prototypes of SARS-CoV due to the calculated binding free energy of SZ3-rat complex being comparable to those of TOR2–civet cat and GD–human complexes. In other words, a probable cross-species transmission chain of SARS-CoV is from previously identified bats (Li et al., 2005a) through rats and subsequently civet cats (Consortium, 2004; Guan et al., 2003) to human, coupled with molecular adaptation to host species-specific factors via high-rate mutations. In fact, the possible vector role of rats in the SARS outbreak at Amoy Gardens has been previously implied (Ng, 2003) and the confirmed first case of human SARS-CoV infection was reported to have only contacts with rats (Liang et al., 2004). In particular, a very recent review (Wang et al., 2006) suggested serologic evidences of rats susceptible to SARS-CoV whereas the detailed information still waits for publication. The viral strain they used might account for the sharp conflict between in vivo and in vitro experimental observations. Certainly, the molecular epidemiological surveillance and subsequent biochemical investigations are clearly necessary to test this hypothesis. As proposed previously (Holmes and Rambaut, 2004), the surveillance should be focused on the high population density wildlife spatially and temporally in close contacts with human community, such as birds, bats and rodents. In addition, the sampling and sequencing of those animal ACE2 in combination with the computational approaches adopted here could be informative for the rapid identification of viral mutants of high likelihood for the utilization of both animal and human receptors, the co-infection of target cells of different host species, and the expansion of host range, or even worse the reemergence of this zoonosis in human population.

In sharp contrast to the predictions based on the complex AE, the calculations derived from the complex BF display an unexpected pattern in conflict with biochemical evidences. For instance, the binding energy of the TOR2–civet cat complex is the minimum among those of all the 12 BF-based structures and even markedly lowers than that of the wild type (TOR2–human) with a difference of more than 3 kcal/mol. However, pseudovirus infections and binding assays (Li et al., 2005b) do not support this indication but find that the TOR2–civet cat complex has an equivalent binding affinity with reference to the TOR2–human complex. Besides, disagreements between computations and experiments are simultaneously found for interactions of human or civet cat receptor with the other two viral RBDs (SZ3 and GD). For each of the two receptors, the SZ3 RBD binds to ACE2 with a lower free energy than the GD RBD, opposite to the rank speculated from biochemical analyses (Li et al., 2005b). Obviously, the energetic effects of viral strain-specific or host species-specific variations on the RBD–ACE2 complex formation under the condition of inhibitor-bound receptor still need further experimental characterization. But in this case, the poor predictability of BF complex prohibits it from further consideration.

Acknowledgements

We thank T. Kortemme, J.J. Gray and G. Cheng for their suggestions on the methodology. We also thank L.F. Wang, H. Field and V. Deubel for helpful discussions.

References

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucleic Acids Res. 28, 235–242.

Consortium, C.S.M.E., 2004. Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. Science 303, 1666–1669.

Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L., Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., Butt, K.M., Wong, K.L., Chan, K.W., Lim, W., Shortridge, K.F., Yuen, K.Y., Peiris, J.S., Poon, L.L., 2003. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science 302, 276–278.

Hofmann, H., Pohlmann, S., 2004. Cellular entry of the SARS coronavirus. Trends Microbiol. 12, 466–472.

Holmes, E.C., Rambaut, A., 2004. Viral evolution and the emergence of SARS coronavirus. Phil. Trans. Roy. Soc. Lond. B: Biol. Sci. 359, 1059–1065.

Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C., Choe, H., Farzan, M., 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426, 450–454.

Li, W., Greenough, T.C., Moore, M.J., Vasilieva, N., Somasundaran, M., Sullivan, J.L., Farzan, M., Choe, H., 2004. Efficient replication of severe acute respiratory syndrome coronavirus in mouse cells is limited by murine angiotensin-converting enzyme 2. J. Virol. 78, 11429–11433.

Li, F., Li, W., Farzan, M., Harrison, S.C., 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 309, 1864–1868.

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., Wang, H., Crameri, G., Hu, Z., Zhang, H., Zhang, J., McEachern, J., Field, H., Daszak, P., Eaton, B.T., Zhang, S., Wang, L.F., 2005a. Bats are natural reservoirs of SARS-like coronaviruses. Science 310, 676–679.

Li, W., Zhang, C., Sui, J., Kuhn, J.H., Moore, M.J., Luo, S., Wong, S.K., Huang, I.C., Xu, K., Vasilieva, N., Murakami, A., He, Y., Marasco, W.A., Guan, Y., Choe, H., Farzan, M., 2005b. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. EMBO J. 24, 1634–1643.

Li, W., Wong, S.K., Li, F., Kuhn, J.H., Huang, I.C., Choe, H., Farzan, M., 2006. Animal origins of the severe acute respiratory syndrome coronavirus: insight from ACE2-S–protein interactions. J. Virol. 80, 4211–4219.

Liang, G., Chen, Q., Xu, J., Liu, Y., Lim, W., Peiris, J.S., Anderson, L.J., Ruan, L., Li, H., Kan, B., Di, B., Cheng, P., Chan, K.H., Erdman, D.D., Gu, S., Yan, X., Liang, W., Zhou, D., Haynes, L., Duan, S., Zhang, X., Zheng, H., Gao, Y., Tong, S., Li, D., Fang, L., Qin, P., Xu, W., 2004. Laboratory diagnosis of four recent sporadic cases of community-acquired SARS, Guangdong Province, China. Emerg. Infect. Dis. 10, 1774–1781.

Liu, Y., Kuhlman, B., 2006. RosettaDesign server for protein design. Nucleic Acids Res. 34, W235–W238.

Marra, M.A., Jones, S.J., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Butterfield, Y.S., Khattra, J., Asano, J.K., Barber, S.A., Chan, S.Y., Cloutier, A., Coughlin, S.M., Freeman, D., Girn, N., Griffith, O.L., Leach, S.R., Mayo, M., McDonald, H., Montgomery, S.B., Pandoh, P.K., Petrescu, A.S., Robertson, A.G., Schein, J.E., Siddiqui, A., Smailus, D.E., Stott, J.M., Yang, G.S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T.F., Bowness, D., Czub, M., Drebot, M., Fernando, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G.A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R.C., Krajden, M., Petric, M., Skowronski, D.M., Upton, C., Roper, R.L., 2003. The genome sequence of the SARS-associated coronavirus. Science 300, 1399–1404.

Ng, S.K., 2003. Possible role of an animal vector in the SARS outbreak at Amoy Gardens. Lancet 362, 570–572.

Rota, P.A., Oberste, M.S., Monroe, S.S., Nix, W.A., Campagnoli, R., Icenogle, J.P., Penaranda, S., Bankamp, B., Maher, K., Chen, M.H., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J.L., Chen, Q., Wang, D., Erdman, D.D., Peret, T.C., Burns, C., Ksiazek, T.G., Rollin, P.E., Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Gunther, S., Osterhaus, A.D., Drosten, C., Pallansch, M.A., Anderson, L.J., Bellini, W.J., 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300, 1394–1399.

Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., Serrano, L., 2005. The FoldX web server: an online force field 10.1093/nar/gki387. Nucleic Acids Res. 33, W382–W388.

Sobolev, V., Eyal, E., Gerzon, S., Potapov, V., Babor, M., Prilusky, J., Edelman, M., 2005. SPACE: a suite of tools for protein structure prediction and analysis based on complementarity and environment 10.1093/nar/gki398. Nucleic Acids Res. 33, W39–W43.

Towler, P., Staker, B., Prasad, S.G., Menon, S., Tang, J., Parsons, T., Ryan, D., Fisher, M., Williams, D., Dales, N.A., Patane, M.A., Pantoliano, M.W., 2004. ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. J. Biol. Chem. 279, 17996–18007.

Wang, L.-F., Shi, Z., Zhang, S., Field, H., Daszak, P., Eaton, B.T., 2006. Review of bats and SARS. Emerg. Infect. Dis..

Wong, S.K., Li, W., Moore, M.J., Choe, H., Farzan, M., 2004. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. J. Biol. Chem. 279, 3197–3201.