PGE_2 receptors, EP_1, EP_3, and EP_4, have Pro at the corresponding site. As described above, the ancestor of the PNRs was considered to be the PGE_2 receptor. The observation of Pro in EP_1, EP_3, and EP_4 at the site suggests that the original residue in the ancestral PGE_2 receptor was Pro and that the ligand recognition mechanism of EP_2 may be different from those of the other PGE_2 receptors and the ancestral receptor, although they share PGE_2 as their ligand. DP lacks binding affinity to PGE_2 [47]. In DP, the residue corresponding to $S28^{1.39}$ is substituted with Gly. Our simulation study of DP suggested that the Gly residue is not involved in the recognition of PGD_2. Kobayashi et al. [46] generated a chimera between IP and DP, in which ICL1, TM2, and ECL1 were derived from IP to the remaining parts of the molecule were from DP. The chimera did not bind to either PGE_2 or PGD_2. When the Gly on TM1 in the DP region of the chimera was substituted with Ser, the mutant acquired the ability to bind PGE_2. The experiment seems to indirectly support the importance of the Ser residue for the recognition of PGE_2.

On the other hand, $K76^{2.54}$ was specifically conserved in DP. Our simulation study suggested that this residue is involved in the recognition of the hydroxy group of the ω chain of PGD_2. No residues interacted with either the hydroxy group or the carbonyl oxygen of the cyclopentane ring of PGD_2 in our simulation. In contrast, no residue that could interact with the hydroxy group of the ω chain of PGE_2 was detected in the simulation of EP_2. Therefore, the hydrogen bond formation between the hydroxy group and $K76^{2.54}$ may regulate the specific ligand recognition by DP. As described above, Li et al. [14] suggested that the Lys can form a hydrogen bond with the carbonyl oxygen in the cyclopentane ring of PGD_2, based on their MD simulation. Thus, our result is inconsistent with their report. However, the average distance between $K76^{2.54}$ and the carbonyl oxygen was 4.9 Å, and some rotamers of the residue were in a position that could interact with the oxygen (data not shown). Therefore, it is still possible that $K76^{2.54}$ also interacts with the carbonyl oxygen. For both PGE_2 and PGD_2, the α chain extended toward the outside of the receptors, whereas the cyclopentane ring and the ω chain extended between TM1 and TM2. The modification of the amino acid residues on the two helices is considered to have been effective to change the ligand recognition mechanism.

The D(E)RY motif is a conserved amino acid stretch in the cytoplasmic region of TM3. The cytoplasmic interaction network between the D(E)RY motif and TM6 is considered to be involved in the activation of the receptor. The network is called an "ionic lock", which is disrupted during the activation to release the constraints on the relative movement of TM3 and TM6 in some class A GPCRs, such as rhodpsin and β_2-adrenergic receptor [48]. The salt bridge between the first and second residues of the motif in the inactivated form is also observed [15]. EP_2 has the canonical D(E)RY motif. $E133^{3.49}$ of the motif did not form a salt bridge with the second residue of the motif, $R134^{3.50}$, but interacted with R146 (ICL2). $R134^{3.50}$ formed a salt bridge with $E259^{6.30}$. To examine the release of the ionic lock by the simulation study, more time steps may be required. In contrast to the canonical motif of EP_2, the D(E)RY motif of DP is mutated. The highly conserved Arg at the second position of the motif is substituted with Cys. Rosenkilde et al. [19] examined 365 human rhodopsin-like GPCRs and reported that only 3% lack the basic residue at this position. In our simulation study, the interaction between the motif and TM6 was absent in DP, due to the mutation. However, DP can exert its signal transduction activity. Therefore, DP may use a different mechanism for receptor activation, corresponding to the ionic lock. This hypothesis seemed to be consistent with the results of the experimental studies on other class A GPCRs. The D(E)RY motifs of various GPCRs have been subjected to amino acid substitution experiments to examine their functional meanings [49]. Such mutants are expected to be constitutively active, since they cannot form the ionic lock. Actually, some GPCR mutants are constitutively active. For example, FP becomes constitutively active by the mutation of the D(E)RY motif [50]. However, quite a few mutants are not constitutively active, such as the mutant of TP [51]. Both FP and TP have the canonical D(E)RY motif. As shown in Fig. 2, EP_2 and DP belong to one of the subtrees, whereas FP and TP are members of the other subtree. Therefore, it seems that the activation mechanism of PNRs may have rapidly diverged during evolution, and at least DP and TP acquired the activation mechanism without the ionic lock. Further experimental and computational studies are needed to solve this problem.

We have described a synergistic study integrating different computational approaches. The evolutionary information obtained from sequence comparisons is useful to select proper targets for computational studies and to identify the candidates of functionally important residues. Structural information obtained from homology modeling and docking analyses provides further clues to refine the candidate residues. Molecular dynamics simulations provided dynamic views for this study, which could not be obtained from static analyses with sequence comparisons and modeling. Thus, combining the information obtained from different computational approaches with that from the literature seems to be more efficient than using an individual approach, if the proper combination is adopted.

In this study, we used the PNRs as a concrete example for the application of the synergistic study. However, many questions about the PNRs still remain. For example, CRTH2 is known to function as a PGD_2 receptor [7], but

the evolutionary origin of the receptor differs from that of DP [8], as described above. Whether CRTH2 and DP share the same ligand recognition mechanism is an interesting subject, not only from a pharmaceutical viewpoint but also from an evolutionary perspective. A synergistic study may provide clues to address this issue. Computational analyses will accelerate the studies of PNRs.

Acknowledgments We thank Drs. Wataru Nemoto, Kentaro Tomii, and Makiko Suwa of CBRC for useful discussions on this work. HD was supported in part by the Global COE program, "an In Silico Medicine", at Osaka University and Grants-in-Aid (Nos. 20650012 and 19650072) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. HT was partially supported by Targeted Proteins Research Program (TPRP).

References

1. Funk CD (2001) Science 294:1871–1875
2. Samuelsson B, Morgenstern R, Jakobsson PJ (2007) Pharmacol Rev 59:207–224
3. Wang MT, Honn KV, Nie D (2007) Cancer Metastasis Rev 26:525–534
4. Schuligoi R et al (2010) Pharmacology 85:372–382
5. Jones RL, Giembycz MA, Woodward DF (2009) Br J Pharmacol 158:104–145
6. Matsuoka T, Narumiya S (2007) Sci World J 7:1329–1347
7. Hirai H et al (2001) J Exp Med 193:255–261
8. Hata AN, Breyer RM (2004) Pharmacol Ther 103:147–166
9. Kedzie KM, Donello JE, Krauss HA, Regan JW, Gil DW (1998) Mol Pharmacol 54:584–590
10. Chang C, Negishi M, Nishigaki N, Ichikawa A (1997) Biochem J 322:597–601
11. Stitham J, Stojanovic A, Merenick BL, O'Hara KA, Hwa J (2002) J Biol Chem 278:4250–4257
12. Funk CD, Furci L, Moran N, Fitzgerald GA (1993) Mol Pharmacol 44:934–939
13. Neuschäfer-Rube F, Engemaier E, Koch S, Böer U, Püschel GP (2003) Biochem J 371:443–449
14. Li Y et al (2007) J Am Chem Soc 129:10720–10731
15. Rosenbaum DM, Rasmussen SG, Kobilka BK (2009) Nature 459:356–363
16. Scheer A, Fanelli F, Costa T, De Benedetti PG, Cotecchia S (1996) EMBO J 15:3566–3578
17. Wess J (1998) Pharmacol Ther 80:231–264
18. Ballesteros J, Palczewski K (2001) Curr Opin Drug Discov Devel 4:561–574
19. Rosenkilde MM, Kledal TN, Schwartz TW (2005) Mol Pharmacol 68:11–19
20. Lu ZL, Curtis CA, Jones PG, Pavia J, Hulme EC (1997) Mol Pharmacol 51:234–241
21. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Nucleic Acids Res 25:3389–3402
22. Ballesteros JA, Weinstein H (1995) Methods Neurosci 25:366–428
23. Katoh K, Misawa K, Kuma K, Miyata T (2002) Nucleic Acids Res 30:3059–3066
24. Katoh K, Toh H (2008) Brief Bioinform 9:286–298
25. Saitou N, Nei M (1987) Mol Biol Evol 4:406–425
26. Felsenstein J (1996) Methods Enzymol 266:418–427
27. Jones DT, Taylor WR, Thornton JM (1992) Comput Appl Biosci 8:275–282
28. Felsenstein J (1985) Evolution 39:783–791
29. Felsenstein J (1993) PHYLIP (phylogeny inference package), version 3.5c. University of Washington, Seattle
30. Adachi J, Hasegawa M (1996) MOLPHY (programs for molecular phylogenetics), version 2.3b3. Institute of Statistical Mathematics, Tokyo
31. Halgren TA (1996) J Comput Chem 17:490–519
32. Labute P (2008) J Comput Chem 29:1693–1698
33. Wiederstein M, Sippl MJ (2007) Nucleic Acids Res 35:W407–W410
34. Sippl MJ (1993) Proteins 17:355–362
35. Goto J, Kataoka R (2008) J Chem Inf Model 48:583–590
36. Bowers KJ, Chow E, Xu H, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kolossvary I, Moraes MA, Sacerdoti FD, Salmon JK, Shan Y, Shaw DE (2006) In: Proceedings of the ACM/IEEE conference on Supercomputing, Tampa, November 11-17, ACM New York, USA. doi:10.1145/1188455.1188544
37. Banks JL, Beard HS, Cao Y, Cho AE, Damm W, Farid B, Felts AK, Halgren TA, Mainz DT, Maple JR, Murphy R, Philipp DM, Repasky MP, Zhang LY, Berne BJ, Friesner RA, Gallicchio E, Levy RM (2005) J Comput Chem 26:1752–1780
38. Krautler V (2001) J Comput Chem 22:501–508
39. Darden T, York D, Pedersen L (1993) J Chem Phys 98:10089–10092
40. Lyne PD, Lamb M, Saeh JC (2006) J Med Chem 49:4805–4808
41. Toh H, Ichikawa A, Narumiya S (1995) FEBS Lett 361:17–21
42. Fritze O et al (2003) Proc Natl Acad Sci USA 100:2290–2295
43. Paila YD, Tiwari S, Chattopadhyay A (2008) Biochim Biophys Acta 1788:295–302
44. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J (2006) Nucleic Acids Res 34:W116–W118
45. Jaakola V-P, Prilusky J, Sussman JL, Goldman A (2005) Protein Eng Des Sel 18:103–110
46. Kobayashi T, Ushikubi F, Narumiya S (2000) J Biol Chem 275:24294–24303
47. Tsuboi K, Sugimoto Y, Ichikawa A (2002) Prostaglandins Other Lipid Mediat 68–69:535–556
48. Vogel R, Mahalingam M, Lüdeke S, Huber T, Siebert F, Sakmar TP (2008) J Mol Biol 380:648–655
49. Rovati GE, Capra V, Neubig RR (2007) Mol Pharmacol 71:959–964
50. Pathe-Neuschäfer-Rube A, Neuschäfer-Rube F, Püschel GP (2005) Biochem J 388:317–324
51. Ambrosio M, Fanelli F, Brocchetti S, Raimondi F, Mauri M, Rovati GE, Capra V (2010) Cell Mol Life Sci 67:2979–2989
52. Tusnády GE, Dosztányi Z, Simon I (2005) Bioinformatics 21:1276–1277

Theor Chem Acc (2011) 130:1197–1202
DOI 10.1007/s00214-011-1048-z

REGULAR ARTICLE

Antigen–antibody interactions of influenza virus hemagglutinin revealed by the fragment molecular orbital calculation

Akio Yoshioka · Kazutomo Takematsu · Ikuo Kurisaki · Kaori Fukuzawa · Yuji Mochizuki · Tatsuya Nakano · Eri Nobusawa · Katsuhisa Nakajima · Shigenori Tanaka

Received: 12 May 2011 / Accepted: 16 September 2011 / Published online: 4 October 2011

Abstract Effective interactions between amino acid residues in antigen–antibody complex of influenza virus hemagglutinin (HA) protein can be evaluated in terms of the inter-fragment interaction energy (IFIE) analysis with the fragment molecular orbital (FMO) method, in which each fragment contains the side chain of corresponding amino acid residue. We have carried out the FMO-MP2 (second-order Moeller–Plesset) calculation for the complex of HA antigen and Fab antibody of influenza virus H3N2 A/Aichi/2/68 and obtained the IFIE values between each amino acid residue in HA and the whole antibody as the sums over the residues contained in the latter. Combining this IFIE data with experimental data for hemadsorption activity of HA mutants, we succeeded in theoretically explaining the mutations in HA observed after the emergence of influenza virus H3N2 A/Aichi/2/68 in an earlier study, except for those of THR83. In the present study, we employ an alternative way of fragment division in the FMO calculation at the carbonyl C site of the peptide bond instead of the C_α site used in the previous work, which provides revised IFIE values consistent with all the historical mutation data in the antigenic region E of HA including the case of THR83 in the present prediction scheme for probable mutations in HA.

Keywords Influenza virus · Hemagglutinin · Antibody · Fragment molecular orbital · Inter-fragment interaction energy

Dedicated to Professor Akira Imamura on the occasion of his 77th birthday and published as part of the Imamura Festschrift Issue.

A. Yoshioka · S. Tanaka (✉)
Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan
e-mail: tanaka2@kobe-u.ac.jp

K. Takematsu · S. Tanaka
Graduate School of Human Development and Environment, Kobe University, 3-11 Tsurukabuto, Nada-ku, Kobe 657-8501, Japan

I. Kurisaki · S. Tanaka
Graduate School of Science and Technology, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan

K. Fukuzawa
Mizuho Information and Research Institute Inc., 2-3 Kanda Nishiki-Cho, Chiyoda-ku, Tokyo 101-8443, Japan

Y. Mochizuki
Department of Chemistry and Research Center for Smart Molecules, Faculty of Science, Rikkyo University, 3-34-1 Nishi-ikebukuro, Toshima-ku, Tokyo 171-8501, Japan

T. Nakano
Division of Medicinal Safety Science, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

E. Nobusawa
National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

K. Nakajima
Department of Virology, Medical School, Nagoya City University, 1 Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya 467-8601, Japan

Abbreviations

HA	Hemagglutinin
FMO	Fragment molecular orbital
IFIE	Inter-fragment interaction energy
MP2	Moeller–Plesset second-order perturbation

1 Introduction

Hemagglutinin (HA), a major antigenic protein of the influenza virus, is a homo-trimeric glycoprotein situated on the viral surface [1]. HA plays an important role in the early stage of infection such as the binding to the receptors (sialic acids) on the host cells and the trigger for the fusion between virus and endosome membranes. The receptor-binding site (RBS) is located on the membrane-distal globular domain of HA. The neutralizing antibody targeting the antigenic regions located around the RBS then prohibits the virus from binding to the receptors. Antigenic variants are generally selected during circulation of the viruses among human population [2]. In these variants (mutants), amino acid differences are observed in the antigenic regions compared to the original viruses. The structural analyses of HA-antibody complexes have provided the information about the amino acid residues on HA directly interacting with the antibody in the complexes. Amino acid changes (mutations) at these positions allow the virus to escape from the neutralizing antibody.

In an earlier study [3], we performed a quantum-chemical, electron-correlated second-order Moeller–Plesset (MP2) perturbation calculation [4] for the HA antigen–antibody system of influenza virus H3N2 A/Aichi/2/68 [5] with the fragment molecular orbital (FMO) method [6, 7], in which a large antigen–antibody biomolecular system was divided into a collection of many fragments corresponding to amino acid residues. On the basis of the calculated inter-fragment interaction energies (IFIEs) [8–14] representing the molecular interactions between the amino acid residues in the antigen–antibody complex, we identified those residues in the antigenic region E [15, 16] of HA protein that were significantly recognized by the Fab fragment of antibody [17, 18] with strongly attractive interactions. Combining these IFIE results with the data of hemadsorption experiments [19, 20] by which the mutation-prohibited sites were specified enabled us to explain most of those mutation sites actually observed (five of six residues) as a benchmark test, which would thus provide a promising method for predicting the HA residues that have a high probability of forthcoming mutation.

In spite of its success, only one residue site in HA has remained to be explained in the earlier study [3], whose mutation cannot be accounted for appropriately in terms of the proposed prediction scheme. That is the THR83 in the antigenic region E of HA, which shows a repulsive interaction with the Fab antibody in the FMO calculation, but has mutated three times [19, 20] after the emergence of the H3N2 influenza virus in 1968. In our prediction method, the amino acid residue that shows an attractive interaction with the antibody was supposed to have a high probability for mutation in order to escape from antibody recognition. One of the possible reasons for the disagreement with the observation can be ascribed to the fragment division method employed in the FMO calculation. We relied on the fragment division at the C_α site in the polypeptide chain according to the usual FMO recipe [6, 7, 21] for reducing the computational error in total energy, while this method of fragmentation may be less natural than the division at the peptide bond in light of biochemical function (see also the Sect. 4 below). Therefore, we here attempt an alternative method of fragment division for the IFIE calculation in the FMO method to overcome this difficulty. Concerning the accuracy of the calculated energies, the fragment division at the atoms other than the C_α site may cause additional energy errors in IFIEs by the order of less than 1 kcal/mol [6, 7, 21]. (Note that the IFIE values calculated with different fragmentations differ mainly due to the difference in the fragment units employed in the FMO calculations.)

The present article is organized as follows. In Sect. 2, the models and methods employed in the present study are illustrated. The calculated results obtained by the two kinds of the fragmentations are shown and compared in Sect. 3. We thus find in Sect. 4 that the IFIEs calculated through the fragmentation at the peptide bond give a description for the antigen–antibody interactions, which is more consistent with the main stream amino acid changes observed in the antigenic region E of HA. Section 5 concludes with a summary.

2 Models and methods

We employ an HA-Fab antigen–antibody system of H3N2 A/Aichi/2/68 influenza virus (PDB ID: 1EO8) [5] in the present analysis. Before performing the FMO-IFIE calculations, we added the missing hydrogen atoms in the complex and optimized their positions with the aid of MMFF94x force field [22] on the MOE (Molecular Operating Environment, Chemical Computing Group Inc.) software. Then, the FMO-MP2/6-31G* [3, 4] and the corresponding classical force-field (Amber ff99 [23]) calculations have been carried out for the antigen–antibody complex with 921 residues and 14,086 atoms, in which the monomer structure of HA [5] is employed. In contrast to

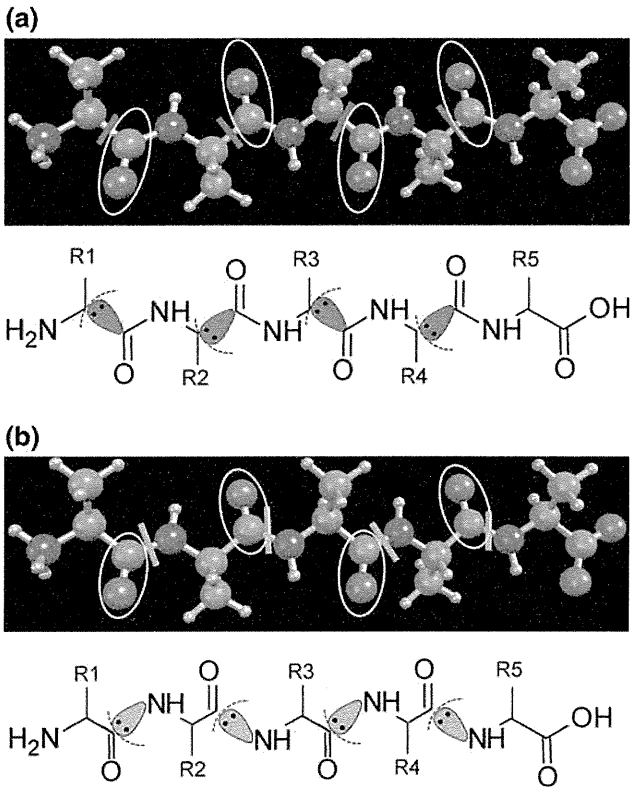

Fig. 1 Two methods of fragment division for the polypeptide chain. **a** The fragment division at the C_α site employed in the usual FMO calculations. **b** The fragment division at the carbonyl C site of the peptide bond employed in the present study. In the *lower panels*, the *dotted line* shows the fragmentation border, where chemical bond (orbital marked in *red* or *blue*) and electrons (represented by *dots*) are partitioned according the conventional FMO recipe [6, 7, 21]

the earlier study, where the fragmentation was performed at the C_α site of the C–C bond according to the usual FMO prescription [6, 7, 21], the fragmentation is performed at the carbonyl C site of the peptide (C–N) bond in the present analysis. The comparison between these two methods of the fragmentation is illustrated in Fig. 1.

We obtained the IFIEs [8–14] on the bases of these force-field and FMO (two-body FMO expansion [21] in this study) calculations and summed the values of IFIEs over the residues contained in the antibody. It is noted here that the IFIE is equivalent to the pair interaction energy (PIE) employed in other FMO studies [21]. These IFIE sums for each antigenic region A–E demonstrate that the amino acid residues in the vicinity of the antigenic region E are significantly recognized by the antibody through strongly attractive interactions, indicating that these residues are highly responsible for the binding affinity between the HA antigen and the Fab fragment of antibody. We therefore focus on 24 residues in the vicinity of the antigenic region E (residue numbers 62–85) in the present analysis.

It is noted, in passing, that we employ in this study an energetically optimized structure of complex on the basis of the PDB (Protein Data Bank) registered structure as a representative snapshot. Structural fluctuations of protein complex in environmental solvent are thus neglected in the present analysis, which may be justified by the consideration that the optimized structure should be observed dominantly by the energetical reason. It is also supposed that the screening effect due to solvent [24] would not change the relative order or the sign of IFIEs. In this regard, the MP2 electron-correlation method [3, 4, 8–14] is employed in the FMO calculation in order to appropriately describe the weak dispersion interactions in addition to other quantum-mechanical effects such as electronic polarization and charge transfer.

3 Calculated results

First, we show the result for the IFIEs between each residue in HA and the Fab fragment of antibody obtained by the classical force-field calculation, in which the fragmentation is performed at the carbonyl C site of the peptide bond. Figure 2a–c demonstrate the calculated results for charged, hydrophobic and polar residues in the antigenic region E, respectively, in comparison with the earlier FMO results to which the fragmentation at the C_α site was applied. As seen in the figures, the qualitative tendency of the interactions is well correlated in both calculations on the whole, while the classical force-field calculation seems to underestimate the contribution associated with the attractive dispersion interaction; the differences for the charged residues may be partially ascribed to the neglects of electronic polarization and charge transfer in the classical calculation. A significant, qualitative difference between the two calculations is then observed for those residues such as GLN80 and THR83, which may be attributed to the difference in the ways of fragment division as mentioned earlier. In particular, the latter residue is important in the present context of mutation prediction. If we would employ the IFIE value between THR83 and antibody obtained through the force-field calculation, we can assign this residue as a probable candidate for forthcoming mutation because the mutation at the THR83 site could diminish the currently attractive interaction with the antibody, which would be favorable for the escape from the antibody pressure. This is then consistent with the observation [19, 20] that the THR83 site has undergone mutations in the actual influenza viruses.

We have next carried out the FMO-MP2/6-31G* calculation for the identical antigen–antibody structure. In this calculation, we have modified the method of fragmentation from that employed in the earlier work, so that five residues including THR83 are separated at the carbonyl C site of the

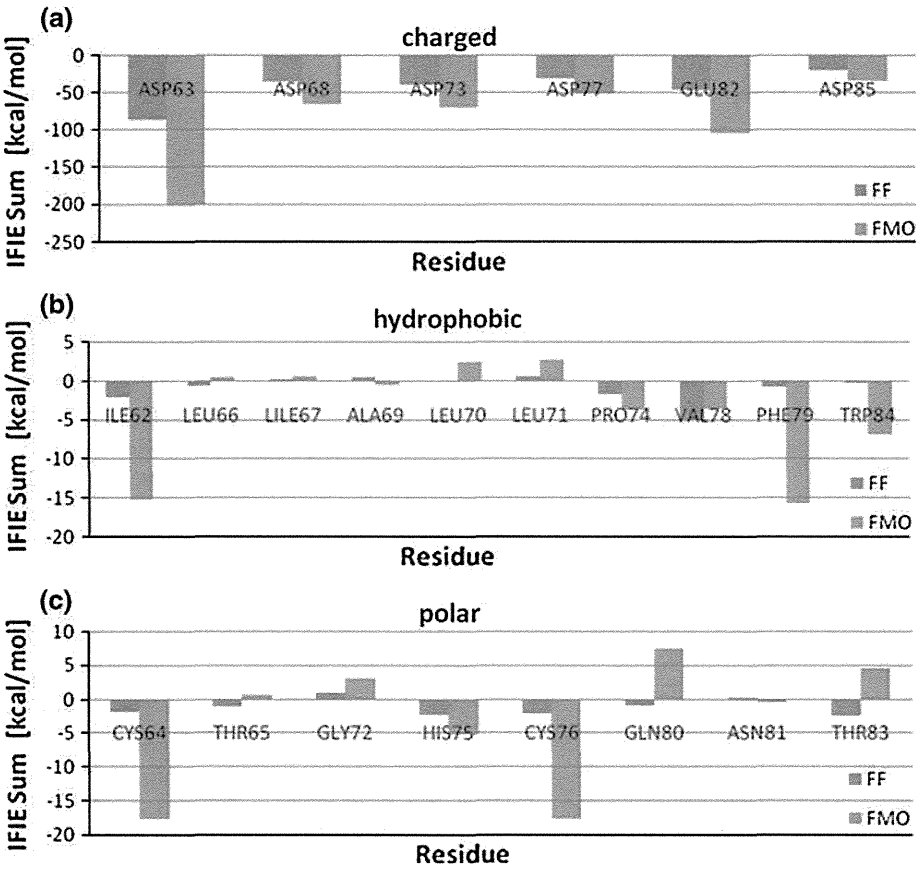

Fig. 2 IFIE values between the whole Fab antibody and each amino acid residue in HA. The *blue* and *red bars* represent the results obtained by the classical force-field (FF) calculation with the present fragmentation of Fig. 1b and by the FMO-MP2/6-31G* calculation with the usual (previous) FMO fragmentation of Fig. 1a, respectively. **a** Charged residues. **b** Hydrophobic residues. **c** Polar residues

peptide bond instead of the C_α site. Table 1 shows the IFIE values between each residue fragment and the antibody, which are compared between the previous and present ways of FMO fragmentation. Most interesting in this table is the IFIE values for THR83 and TRP84. The IFIE for THR83 changes from 4.52 kcal/mol in the previous scheme to −4.93 kcal/mol in the present scheme, which is in line with the result observed in the classical force-field calculation above; on the other hand, that for TRP84 changes from −6.85 to 1.25 kcal/mol. These results can be accounted for as consequences of the modification of the fragmentation, as will be addressed in the following section, and provide a consequence consistent with the observation [19, 20] for mutations of HA in the H3N2 influenza virus.

4 Discussion

As addressed in Sect. 2, the fragment assignment of the carbonyl group C=O in the main chain of polypeptide is different between the two types of fragmentations shown in Fig. 1; the fragmentation was performed at the C_α site in the usual FMO scheme, while it is performed at the carbon site of the carbonyl group C=O constituting the peptide bond in the scheme proposed in the present study. When we consider the IFIE between THR83 in HA and ARG98 in Fab antibody, this difference causes a significant consequence. Figure 3 illustrates the configuration of THR83 and ARG98 in the antigen–antibody complex. The carbonyl group next to the side chain of THR83 has an attractive interaction with ARG98. In the case of original FMO fragmentation, this carbonyl group belongs to the fragment containing the side chain of TRP84, then bringing about an attractive IFIE between the TRP84 fragment and the antibody (see −6.85 kcal/mol in Table 1). In the present fragmentation, on the other hand, this carbonyl group belongs to the fragment containing the side chain of THR83. Thus, the IFIE value between the THR83 fragment and the antibody becomes negative (−4.93 kcal/mol) in the present FMO calculation, as seen in Table 1, providing a calculated result consistent with the mutation data in terms of the present prediction scheme.

It is noted that the mutations at all six residue sites, which have been observed in the antigenic region E of HA of H3N2 A/Aichi/2/68, can thus be accounted for by the prediction method based on the present fragmentation scheme. This example would suggest that the fragmentation at the carbonyl site of the peptide bond may be a more natural way of fragment division in FMO calculation in

Table 1 IFIE values (in units of kcal/mol) between the Fab antibody and amino acid residues in the vicinity of THR83

Fragment	IFIE sum (kcal/mol)	Number of atoms
(a) Previous		
GLN80	7.47	17
ASN81	−0.45	14
GLU82	−102.51	15
THR83	4.52	14
TRP84	−6.85	24
ASP85	−34.93	12
LEU86	2.21	19
(b) Present		
GLN80	7.43	17
ASN81	0.63	16
GLU82	−101.98	15
THR83	−4.93	14
TRP84	1.25	24
ASP85	−37.41	10
LEU86	2.23	19

(a) The results obtained by FMO-MP2/6-31G* calculation employing the usual fragment division at the C_α site of the C–C bond shown in Fig. 1a. (b) The results obtained by FMO-MP2/6-31G* calculation employing the present fragment division at the carbonyl C site of the peptide (C–N) bond shown in Fig. 1b. The numbers of atoms contained in each fragment corresponding to amino acid residue (side chain) are also shown

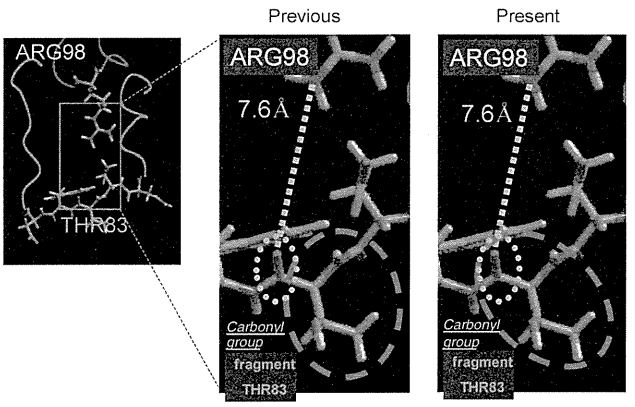

Fig. 3 Configuration of THR83 in HA antigen and ARG98 in Fab antibody. The side chain of THR83 and the carbonyl group interacting with ARG98 are *circled* by the *dashed red line* and the *dotted yellow line*, respectively. These two parts are contained in an identical fragment (THR83) only in the present fragmentation scheme

order to discuss biochemical consequences of mutations, while the original method of fragment division was employed to reduce the computational errors due to the fragmentation as much as possible in the FMO calculations [6, 7, 21]. In this connection, we suppose that the change of the side chain of amino acid residue would modify the local structure of neighboring region of the main chain as well, thus leading to the change in the interaction associated with the main chain. Thus, the effects of the mutations are indirect when the interactions associated with the main chain are important.

For more realistic analysis, we may employ the trimer structure [1] of HA antigen complexed with the Fab dimer antibody. The FMO calculations for this HA trimer system has already been performed [25] at the MP2 and MP3 levels. The prediction of mutations using this FMO-IFIE result and its comparison with that by HA monomer would be interesting, which will be reported elsewhere [26]. It is also remarked that the effect of fluctuating protein structures in aqueous solution on the antigen–antibody interaction [27], which is beyond the scope of the present study, is an issue to be investigated in the future.

5 Conclusion

Our prediction scheme for probable mutations in HA protein of influenza virus depends on the evaluation of interactions between amino acid residues in HA and associated antibody. If the interaction is attractive, the corresponding residue would have a higher probability of forthcoming mutation in order to mitigate the molecular recognition by the antibody. These interaction energies could be evaluated in terms of the IFIEs obtained through the FMO calculation for the antigen–antibody complex. Then, the method of the fragment division employed in the FMO calculation would be important in order to appropriately assign the local interactions associated with pertinent groups in amino acid residues to each fragment. It is noted that these fragmentations are somewhat empirical and lead to the arbitrariness in the IFIE values. In our previous analysis [3] on an HA antigen–antibody system, we relied on the conventional fragmentation scheme in the FMO calculation and found that the prediction scheme worked well for explaining five of the six past mutations in the HA of H3N2 A/Aichi/2/68 influenza virus, but leaving the mutations of THR83 unexplained. In the present study, on the other hand, we have performed a novel fragmentation at the carbonyl C site of the peptide bond, which is more consistent with the biochemical importance of the peptide bond moiety in protein structures. While this fragmentation may bring about increased errors of calculated energies in the FMO approximation, we have improved the agreement between the theoretical prediction and the past mutations observed in the antigenic region E of HA of H3N2 A/Aichi/2/68 influenza virus. This finding may thus suggest various options [28] for us to choose the method of fragmentation employed in the FMO calculation according to the purpose of analysis.

Acknowledgments This work was partially supported by the CREST project of Japan Science and Technology Agency (JST) and by the Health and Labour Sciences Research Grants on Emerging and Re-emerging Infectious Diseases (No. H22-Shinko-Ippan-006) from the Ministry of Health, Labour and Welfare of Japan.

References

1. Barbey-Martin C, Gigant B, Bizebard T, Calder LJ, Wharton SA, Skehel JJ, Knossow M (2002) Virol 294:70
2. Russell CA et al (2008) Science 320:340
3. Takematsu K, Fukuzawa K, Omagari K, Nakajima S, Nakajima K, Mochizuki Y, Nakano T, Watanabe H, Tanaka S (2009) J Phys Chem B 113:4991
4. Mochizuki Y, Yamashita K, Murase T, Nakano T, Fukuzawa K, Takematsu K, Watanabe H, Tanaka S (2008) Chem Phys Lett 457:396
5. Fleury D, Daniels RS, Skehel JJ, Knossow M, Bizebard T (2000) Struct Funct Gen 40:572
6. Kitaura K, Ikeo E, Asada T, Nakano T, Uebayasi M (1999) Chem Phys Lett 313:701
7. Nakano T, Kaminuma T, Sato T, Akiyama Y, Uebayasi M, Kitaura K (2000) Chem Phys Lett 318:614
8. Fukuzawa K, Komeiji Y, Mochizuki Y, Kato A, Nakano T, Tanaka S (2006) J Comput Chem 27:948
9. Fukuzawa K, Mochizuki Y, Tanaka S, Kitaura K, Nakano T (2006) J Phys Chem B 110:16102
10. Kurisaki I, Fukuzawa K, Komeiji Y, Mochizuki Y, Nakano T, Imada J, Chmielewski A, Rothstein SM, Watanabe H, Tanaka S (2007) Biophys Chem 130:1
11. Ito M, Fukuzawa K, Mochizuki Y, Nakano T, Tanaka S (2007) J Phys Chem B 111:3525
12. Ito M, Fukuzawa K, Mochizuki Y, Nakano T, Tanaka S (2008) J Phys Chem A 112:1986
13. Ito M, Fukuzawa K, Ishikawa T, Mochizuki Y, Nakano T, Tanaka S (2008) J Phys Chem B 112:12081
14. Iwata T, Fukuzawa K, Nakajima K, Aida-Hyugaji S, Mochizuki Y, Watanabe H, Tanaka S (2008) Comput Biol Chem 32:198
15. Wiley DC, Wilson IA, Skehel JJ (1981) Nature 289:373
16. Skehel JJ, Stevens DJ, Daniels RS, Douglas AR, Knossow M, Wilson IA, Wiley DC (1984) Proc Natl Acad Sci USA 81:1779
17. Bizebard T, Mauguen Y, Petek F, Rigolet P, Skehel JJ, Knossow M (1990) J Mol Biol 216:513
18. Gigant B, Fleury D, Bizebard T, Skehel JJ, Knossow M (1995) Proteins 23:115
19. Nakajima K, Nobusawa E, Tonegawa K, Nakajima S (2003) J Virol 77:10088
20. Nakajima K, Nobusawa E, Nagy A, Nakajima S (2005) J Virol 79:6472
21. Fedorov DG, Kitaura K (2009) The fragment molecular orbital method. CRC Press, Boca Raton
22. Halgren TA (1996) J Comput Chem 17:490
23. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) J Comput Chem 26:1668
24. Watanabe H, Okiyama Y, Nakano T, Tanaka S (2010) Chem Phys Lett 500:116
25. Mochizuki Y, Yamashita K, Fukuzawa K, Takematsu K, Watanabe H, Taguchi N, Okiyama Y, Tsuboi M, Nakano T, Tanaka S (2010) Chem Phys Lett 493:346
26. Yoshioka A, Fukuzawa K, Mochizuki Y, Yamashita K, Nakano T, Okiyama Y, Nobusawa E, Nakajima K, Tanaka S (2011) J Mol Graph Model 30:110
27. Zhou R, Das P, Royyuru AK (2008) J Phys Chem B 112:15813
28. Tamura K, Inadomi Y, Nagashima U (2007) Bull Chem Soc Jpn 80:721

Journal of Molecular Graphics and Modelling 30 (2011) 110–119

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Prediction of probable mutations in influenza virus hemagglutinin protein based on large-scale ab initio fragment molecular orbital calculations

Akio Yoshioka[a], Kaori Fukuzawa[b], Yuji Mochizuki[c], Katsumi Yamashita[d], Tatsuya Nakano[e], Yoshio Okiyama[f], Eri Nobusawa[g], Katsuhisa Nakajima[h], Shigenori Tanaka[a,*]

[a] Graduate School of System Infrormatics, Kobe University, 1-1, Rokkodai, Nada-ku, Kobe 657-8501, Japan
[b] Mizuho Information and Research Institute, Inc., 2-3, Kanda Nishi-cho, Chiyoda-ku, Tokyo 101-8443, Japan
[c] Department of Chemistry and Research Center for Smart Molecules, Faculty of Science, Rikkyo University, 3-34-1, Nishi-Ikebukuro, Toshima-ku, Tokyo 171-8501, Japan
[d] 1st Manufacturing Industries Solutions Division, NEC Soft Ltd., 1-18-6, Shinkiba, Koto-ku, Tokyo 136-8608, Japan
[e] Division of Safety Information on Drug, Food and Chemicals, National Institute of Health Sciences, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan
[f] Institute of Industrial Science, The University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, Japan
[g] National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan
[h] Department of Virology, Medical School, Nagoya City University, 1, Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya 467-8601, Japan

ARTICLE INFO

Article history:
Received 22 April 2011
Received in revised form 24 June 2011
Accepted 27 June 2011
Available online 6 July 2011

Keywords:
Influenza
Hemagglutinin (HA)
Antibody
Mutation
Fragment molecular orbital (FMO) method

ABSTRACT

Ab initio electronic-state calculations for influenza virus hemagglutinin (HA) trimer complexed with Fab antibody were performed on the basis of the fragment molecular orbital (FMO) method at the second and third-order Møller–Plesset (MP2 and MP3) perturbation levels. For the protein complex containing 2351 residues and 36,160 atoms, the inter-fragment interaction energies (IFIEs) were evaluated to illustrate the effective interactions between all the pairs of amino acid residues. By analyzing the calculated data on the IFIEs, we first discussed the interactions and their fluctuations between multiple domains contained in the trimer complex. Next, by combining the IFIE data between the Fab antibody and each residue in the HA antigen with experimental data on the hemadsorption activity of HA mutants, we proposed a protocol to predict probable mutations in HA. The proposed protocol based on the FMO-MP2.5 calculation can explain the historical facts concerning the actual mutations after the emergence of A/Hong Kong/1/68 influenza virus with subtype H3N2, and thus provides a useful methodology to enumerate those residue sites likely to mutate in the future.

1. Introduction

Hemagglutinin (HA), a major antigenic protein of the influenza virus, is a homo-trimeric glycoprotein. HA plays an important role in the early stage of infection; binding to the receptors (sialic acids) on the host cells and trigger for the fusion between virus and endosome membranes. The receptor binding site (RBS) locates on the membrane-distal globular domain of HA. The neutralizing antibody targeting antigenic regions located around RBS prohibits the virus from binding to the receptors [1,2]. Antigenic variants are generally selected during circulation of the viruses among human population. In these variants, amino acid differences are observed in the antigenic region compared to the original viruses [3]. The structural analyses of HA–antibody complexes have provided the information about the amino acid residues on HA directly interacting with the antibody in the complexes. Amino acid substitutions at these sites allow the virus to escape from the neutralizing antibody [4–6].

In this paper, we first specify the sites of the antigenic region under the antibody pressure based on the X-ray structure of complex of the HA A/Hong Kong/1/68 (H3N2) with the antibody HC63 [7]. This structure is a complex of HA trimer and fragment of antigen binding (Fab) dimer. The pandemic of human influenza in 1968 was caused by the viruses that were human/avian virus reassortants [8]. Moreover, this virus often has undergone antigenic drifts after this outbreak [9]. We may further be faced with influenza pandemics by the causes of antigenic shift with the highly lethal viruses and antigenic drift in the future.

To predict the probable mutation sites of the antigenic region in HA, in an earlier work [10], we developed a practical scheme by combining the inter-fragment interaction energy (IFIE) analysis of fragment molecular orbital (FMO) method with experimental information about the hemadsorption activity of mutants [11,12], in which an HA monomer structure [6] was used. However, it was indicated [7] that the Fab monomer can effectively recognize not only directly linked (e.g., Fab(I) versus HA(I) or Fab(II) versus HA(II))

* Corresponding author. Tel.: +81 78 803 6620; fax: +81 78 803 6621.
E-mail address: tanaka2@kobe-u.ac.jp (S. Tanaka).

1093-3263/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2011.06.011

domains but also cross-linked (e.g., Fab(I) versus HA(II) or Fab(II) versus HA(III)) domains in the complex structure of HA trimer and Fab dimer (see Fig. 1). Thus, we here investigate the interactions of the HA trimer with the Fab dimer by ab initio FMO calculations to quantify the interaction energies at antigenic sites, which would provide a more quantitative measure of how each amino-acid residue is recognized by the Fab antibody.

In the following, we first explain the computational procedure of FMO method in Section 2. Next we show in Section 3 the results for the assessment of the accuracy of calculation, the correlation between hydrogen bonds and IFIE values on the surface of epitopes, and the prediction of amino acid mutations in HA. Concluding remarks are given in Section 4, where we address a possible contribution of the present scheme to the development of influenza vaccine.

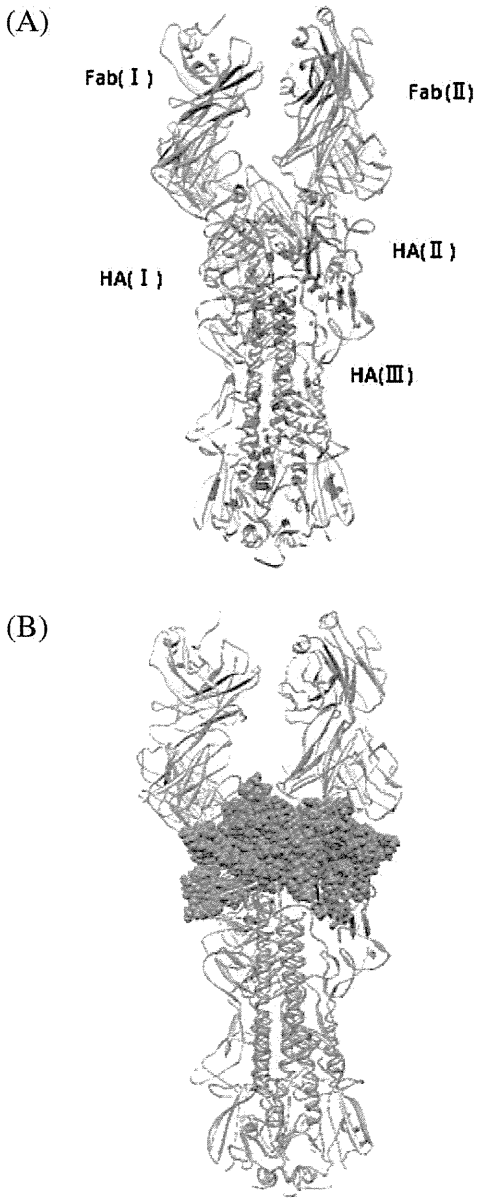

Fig. 1. (A) Visualization of IFIEs between HA trimer (I, II, III) and Fab dimer (I, II) calculated at the FMO-MP2.5/6-31G level. The color represents the sign and strength of the interactions between each residue in the HA trimer and the Fab dimer. For the Fab domain indicated in yellow, the red and blue fragments refer to stabilized and destabilized interactions, respectively, and the deepness of the hue indicates the strength of the interaction. (B) Visualization of antigenic regions A (pink) and B (light blue) by sphere representation. The illustration was created with BioStation Viewer [16]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2. Materials and computational methods

We consider a complex of HA trimer and Fab dimer that is registered in Protein Data Bank (PDB) with ID code of 1KEN [7]. Each HA monomer has the antigenic regions A–E [1,3], although they are further subdivided and sometimes overlaps of these areas have been noted [13,14]. In particular, the Fab has been observed to be bound with the antigenic regions of A, B and E in electron microscope experiments [15]. The neutralization by antibodies is effective for these regions because they are considered to bind to glycoprotein before it binds to sugar moieties of host cell. In the case of 1KEN, it is observed that the Fab dimer is located over both the antigenic regions of A and B (see Fig. 1 [16]). The antibody HC63 [17] also recognizes the multiple HA monomers primarily through the hydrogen bonds, then suggesting the locations of epitopes containing the residues of numbers 136, 137, 153, 158, 159, 186, 187, 189, 190, 192, 193, 225, 226 (direct-linked), and 126, 128, 162, 163, 165 (cross-linked) [7].

The FMO calculations [18,19] have been performed even for enormous proteins like the complex of HA trimer and Fab dimer [20]. We here employ the Hartree–Fock (HF), the second and third-order Møller–Plesset (MP2 and MP3) perturbation methods for comparison. While the HF method is a rough mean-field approximation, the MP2 and MP3 perturbation methods are employed for the description of electron correlations in the present study, where the latter is expected to correct the tendency of overestimation of stabilization energy in the former [21], but too much. Thus, we can accurately evaluate the binding energy in terms of MP2.5 method which provides a half-and-half mixture of MP2 and MP3 energies [20,21].

It is remarked that the effective fragment–fragment interactions in the FMO scheme [22] are obtained in terms of the inter-fragment interaction energy (IFIE) that is defined as

$$\Delta E_{ij} = (E'_{ij} - E'_i - E'_j) + Tr(\Delta P_{ij} V_{ij}), \tag{1}$$

where ΔP_{ij} is a difference density matrix, V_{ij} is an environmental electrostatic potential for fragment dimer ij from other fragments, and E'_i and E'_{ij} are energies of fragment monomer i and dimer ij without environmental electrostatic potential, respectively. These values ΔE_{ij} then represent the interaction energies of an amino acid residue with a ligand or between amino acid residues because each amino acid is assigned as a single fragment in the present analysis [23–25]. The IFIEs were calculated in this study to analyze the interaction pattern and to estimate the contribution of each residue to binding, as seen in Fig. 1.

It is also convenient to introduce [10,26]

$$\Delta E_{ij}^{\text{total}} = \Sigma_j \Delta E_{ij}, \tag{2}$$

which refers to the contribution of each fragment i to the binding affinity with the domain J containing the grouped residues j. It is noted here that

$$\Delta E_{IJ}^{\text{total}} = \Sigma_i \Delta E_{ij}^{\text{total}} \tag{3}$$

represents the inter–domain interaction between the domain I containing the residues i and the domain J containing the residues j.

The structure of the complex of HA antigen and Fab antibody for the FMO calculations was prepared as follows. Starting with the PDB structure (1KEN), the missing hydrogen atoms were complemented by the MOE (Molecular Operating Environment) software (Chemical Computing Group Inc.). The locations of the hydrogen atoms

were then optimized with the MMFF94x force field. The total numbers of residues and atoms were 2351 and 36,160, respectively. We carried out the FMO-HF (Hartree–Fock), MP2 and MP3 calculations [20] with the basis set 6-31G on the Earth Simulator in Yokohama, as in the previous work [27].

3. Results and discussion

3.1. Analysis of inter-domain interaction energies

We first study the inter-domain interactions in the complex of HA trimer and Fab dimer. As seen in Table 1, the calculated value of the interaction energy between the HA trimer and the Fab dimer is +38.1 kcal/mol at the HF level, while the MP2 and MP3 results have been found to be −163.5 kcal/mol and −127.0 kcal/mol, respectively. The reason why the interaction energy in the HF approximation is positive (repulsive) even though the HA trimer should be attractive to the Fab is that the dispersion energies are not appropriately described. In earlier studies [24,28] with the MP2 method for receptor–ligand systems, it has been demonstrated that the dispersion energies play an important role for describing the interactions of biomacromolecules. We evaluate the interaction energy as −145.3 kcal/mol by the MP2.5 method, which would provide a dependable value. The interaction energies between the HA monomers included in the complex are −1445.5 kcal/mol to −1223.1 kcal/mol in the MP2.5, showing the strong binding. The interaction energies between the Fab monomer and the associated HA monomer for Fab(I)–HA(I) and Fab(II)–HA(II) are −359.9 kcal/mol and −372.0 kcal/mol in the MP2.5, respectively. However, the interactions between the Fab monomer and the other HA monomer, and those between the Fab domains were found to be repulsive. These results suggest that the Fab monomers interact attractively only with the bonding HA monomers. It is then noted that, because of the disulfide bridge at the Fab connection site (not seen explicitly in the structure), there is some affinity instead of steric hindrance in this structure in spite of the repulsive interaction [7].

Table 1
IFIE results (in units of kcal/mol) for inter-domain interactions in the complex of HA-trimer (I, II and III) and Fab-dimer (I and II). The calculated values by the HF, MP2, MP3 and MP2.5 methods with the basis set of 6-31G are shown.

Inter-domain	HF	MP2	MP3	MP2.5
Fab dimer–HA trimer	38.1	−163.5	−127.0	−145.3
Fab(I)–HA(I)	−288.8	−367.0	−352.8	−359.9
Fab(I)–HA(II)	177.5	155.5	144.5	150.0
Fab(I)–HA(III)	134.3	134.3	134.3	134.3
Fab(II)–HA(I)	137.0	137.0	137.0	137.0
Fab(II)–HA(II)	−292.7	−380.4	−363.7	−372.0
Fab(II)–HA(III)	170.8	157.0	159.5	158.2
HA(I)–HA(II)	−1022.4	−1280.4	−1237.1	−1258.7
HA(II)–HA(III)	−981.7	−1245.7	−1200.6	−1223.1
HA(I)–HA(III)	−1189.0	−1469.7	−1421.3	−1445.5
Fab(I)–Fab(II)	210.8	197.7	199.5	198.6
Fab dimer–HA(I)	−151.8	−230.0	−215.8	−222.9
Fab dimer–HA(II)	−115.3	−224.9	−205.0	−214.9
Fab dimer–HA(III)	305.1	291.3	293.8	292.6

IFIE value (kcal/mol).

Table 2
Interaction energies (in units of kcal/mol) between residues calculated with FMO-MP2.5/6-31G method and the donor–acceptor distance (Å) of hydrogen bonds in epitopes. The signatures (A)–(J) are associated with those in Fig. 2. As an exceptional case in the calculation, LEU164 is listed instead of VAL163 in (J) because of the fragmentation at C_α employed in the FMO calculation.

Fab	HA (kcal/mol)	Distance (Å)	Fab	HA (kcal/mol)	Distance (Å)
(A)	ASN137		(B)	ASN137	
ASP104	−22.7	1.39	ASP104	−34.54	1.54
			SER32	−8.70	1.98
	TRP153			TRP153	
ASP104	−4.07		ASP104	−1.86	
(C)	GLY158		(D)	GLY158	
PRO60	−0.42		GLY58	0.88	
ALA61	−0.54		VAL59	0.56	
	SER159			SER159	
PRO60	1.36		GLY58	−5.70	1.86
ALA61	−0.43		VAL59	−11.66	1.80
(E)	SER186		(F)	SER186	
TYR100	−8.27	1.80	TYR100	−5.16	2.51
	THR187			THR187	
TYR100	−2.48		TYR100	−3.02	
	GLN189			GLN189	
TYR109	−9.71	1.52	TYR109	−11.12	1.76
	THR192			THR192	
SER57	−11.11	1.76	SER57	0.73	
	SER193			SER193	
SER57	−0.35		SER57	−15.31	1.60
(G)	THR126		(H)	THR126	
SER31	−1.26		SER31	−0.75	
	THR128			THR128	
THR74	0.14		SER28	−2.89	2.96
SER75	−12.70	1.76			
(I)	PR0162		(J)	PR0162	
GLY26	0.23		SER28	−2.27	
	VAL163			LEU 164	
GLY26	0.63		SER31	−11.44	1.54
TYR27	−11.06	2.03	GLY32	−2.08	
	ASN165			ASN165	
SER31	−0.60		SER31	−6.47	
GLY32	−12.55	1.64	GLY32	−12.05	1.52

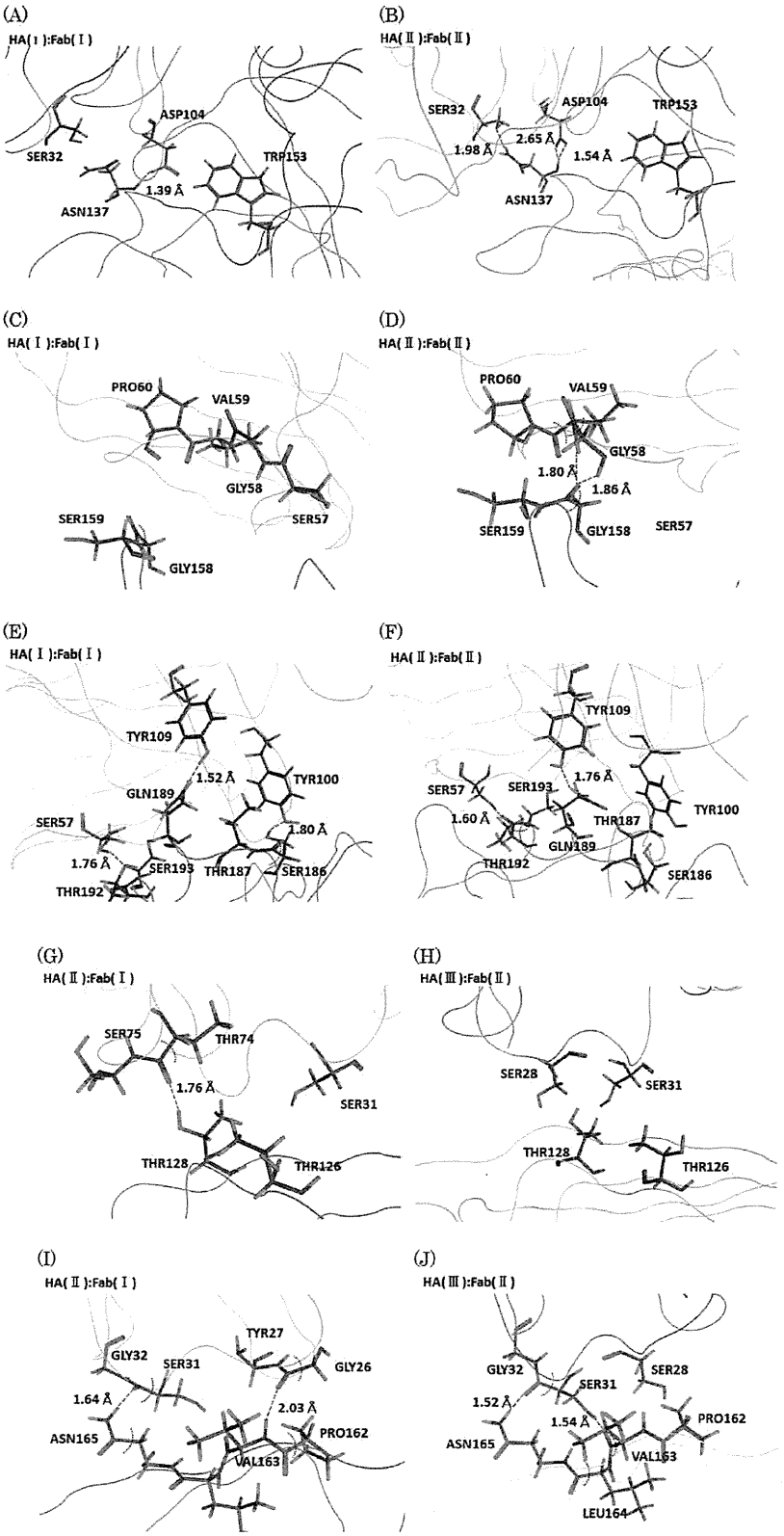

Fig. 2. Details of hydrogen bonds between the Fab and the epitopes in HA. (A)–(F) and (G)–(J) refer to direct-linked and cross-linked domains, respectively; (A), (C), (E) and (B), (D), (F) refer to the interactions between HA(I) and Fab(I) or HA(II) and Fab(II), respectively; Fab in gray, HA(I) in red, HA(II) in blue and HA(III) in yellow. Hydrogen bonds are represented by dotted lines with the distance between donor and acceptor atoms.

Table 3
The hemadsorption activity (p: positive/n: negative; – means no experimental data) and IFIE results (kcal/mol) for the Fab dimer calculated with FMO-MP2.5/6-31G method at the residue sites in antigenic regions A and B of HA. Trimer is the complex of HA(I)–(III), and the mutation years and the types of amino acid after mutation are also listed. The meanings of attached symbols for p sites are as follows. *: Attractive and already mutated site, **: attractive and yet to-be-mutated site, †: repulsive but already mutated site, ††: repulsive and not mutated site.

Region A	p/n	HA(I)	HA(II)	HA(III)	Trimer		Mutation year	Mutation
PHE120	n	0.05	0.39	−0.19	0.24			
ILE121	p	−0.14	−0.76	−0.01	−0.91	*	1995	I-N
THR122	p	0.43	1.24	0.31	1.98	†	1971	T-N
GLU123	–	−46.44	−62.48	−62.62	−171.54			
GLY124	p	0.15	0.24	−0.21	0.19	†	1986, 1996	G-D-S
PHE125	p	−0.72	−1.29	−0.51	−2.52	**		
THR126	p	0.70	−2.74	−0.89	−2.93	*	1973	T-N
TRP127	n	−0.87	−1.15	0.82	−1.20			
THR128	p	0.46	−15.66	−10.86	−26.05	**		
GLY129	n	−0.49	−2.91	−0.39	−3.80			
VAL130	n	0.23	−1.03	0.74	−0.06			
THR131	p	−1.92	−4.16	−1.27	−7.35	*	1987, 2001	T-A-T
GLN132	–	3.02	6.54	4.49	14.05			
ASN133	p	0.86	0.72	1.24	2.81	†	1977, 1991	N-S-D
GLY134	n	−0.13	0.89	−0.20	0.55			
GLY135	p	1.49	−0.03	1.35	2.81	†	1993	G-T
SER136	n	−15.88	−22.56	−0.13	−38.57			
ASN137	p	−24.48	−39.49	−0.54	−64.51	*	1976, 1997	N-Y-S
ALA138	–	−30.89	−31.00	0.89	−61.00			
CYS139	n	−23.09	−19.69	0.11	−42.67			
LYS140	n	21.08	29.32	55.39	105.79		2006	G-I
ARG141	–	33.60	52.72	59.95	146.28			
GLY142	p	−1.65	−3.12	−0.01	−4.78	*	1995	G-R
PR0143	p	−0.43	1.67	0.28	1.52	†	1977	P-S
GLY144	p	−4.71	−3.79	−0.48	−8.97	*	1970, 1982, 1985, 1995	G-D-N-V-R
SER145	p	1.52	−0.65	0.51	1.38	†	1975, 1993, 2004	S-N-K-N
GLY146	p	−6.34	−7.00	−0.48	−13.82	*	1977	G-S
PHE147	p	9.40	7.89	−0.34	16.94	††		
PHE148	n	−3.34	−1.86	0.36	−4.83			
SER149	–	−0.97	−1.42	−1.25	−3.64			
ARG150	–	40.23	55.59	56.94	152.76			
LEU151	n	−0.97	−0.64	−0.46	−2.07			
ASN152	n	3.06	1.05	−0.80	3.31			
TRP153	n	−3.57	−0.17	0.89	−2.84			

Region B	p/n	HA(I)	HA(II)	HA(III)	Trimer		Mutation year	Mutation
LEU154	–	−0.18	−1.72	−1.18	−3.09			
THR155	p	−1.78	−0.29	−0.55	−2.62	*	1971, 1986, 2001	T-Y-H-T
LYS156	p	−24.52	−33.47	82.77	24.78	†	1992, 1995, 2001	E-K-Q-H
SER157	p	0.97	1.02	1.15	3.14	†	1993	S-L
GLY158	p	−1.68	−9.35	−1.47	−12.50	*	1976, 1995	G-E-K
SER159	p	−6.54	−28.09	2.12	−32.52	*	1985, 2002	S-Y-F
THR160	p	3.90	−3.96	0.32	0.25	†	1977	T-K
TYR161	–	1.89	6.83	3.24	11.96			
PR0162	–	−2.61	−3.64	−2.38	−8.64			
VAL163	p	0.93	−4.04	−0.13	−3.24	*	1983	V-A
LEU164	–	0.01	−11.25	−16.43	−27.66			
ASN165	p	−1.16	−2.44	−8.93	−12.53	**		
VAL166	–	0.05	−8.92	−2.13	−11.00			
THR167	–	−0.04	5.16	2.79	7.91			
MET168	–	0.05	−3.88	−3.47	−7.30			
PR0169	–	−0.03	2.79	2.24	5.00			
ASN170	n	0.46	−2.03	−1.61	−3.18			
ASN171	–	−0.29	2.19	1.56	3.46			
ASP172	p	−45.49	−60.16	−59.56	−165.21	*	1976, 1993, 1998	D-G-D-E
ASN173	p	−0.24	0.27	0.04	0.07	†	1982	N-K
PHE174	p	0.32	0.75	0.78	1.85	††		
ASP175	–	−46.03	−49.56	−51.96	−147.55			
LYS176	–	46.24	58.73	59.82	164.80			
LEU177	–	−0.21	1.15	0.72	1.67			
TYR178	n	0.02	−0.76	−0.63	−1.37			
ILE179	n	−0.23	1.01	0.89	1.67			
TRP180	–	0.30	−1.31	−1.86	−2.87			
GLY181	n	0.23	1.07	1.24	2.54			
ILE182	p	−1.04	−1.24	−1.65	−3.94	*	1968	I-V
HIS183	n	2.45	35.21	64.98	102.64			
HIS184	–	33.53	43.09	55.72	132.34			
PR0185	n	6.73	4.00	0.58	11.31			
SER186	–	−15.94	−11.42	0.45	−26.91		1999, 2009	S-G
THR187	–	−12.59	−6.99	−0.02	−19.61			
ASN188	p	5.92	0.53	−1.49	4.96	†	1970	N-D
GLN189	p	−20.65	−28.95	−1.02	−50.61	*	1973, 1987, 1992, 2002	Q-K-R-S-N

Table 3 (*Continued*)

Region B	p/n	HA(I)	HA(II)	HA(III)	Trimer		Mutation year	Mutation
GLU190	–	−11.47	−25.83	−64.32	−101.62		1991	E-D
GLN191	n	5.85	3.99	0.86	10.70			
THR192	p	−12.54	−0.15	−1.35	−14.03	*	1998	T-I
SER193	p	−13.02	−33.20	−3.15	−49.37	*	1972, 1989, 2004	S-N-S-F
LEU194	–	11.86	13.37	−2.04	23.19			
TYR195	–	−0.48	−2.94	−0.50	−3.93			
VAL196	p	−3.64	−6.97	−2.05	−12.66	**		
GLN197	p	−3.37	0.21	1.97	−1.18	*	1977, 1993	Q-R-Q
ALA198	p	1.47	3.85	2.70	8.02	††		
SER199	–	0.67	2.44	1.72	4.82			
GLY200	–	−0.12	−1.41	−0.96	−2.49			
ARG201	–	55.63	55.57	64.96	176.16			
VAL202	n	0.25	−1.09	−1.57	−2.42		2001	V-I
THR203	–	−0.10	1.56	1.55	3.01			
VAL204	n	0.18	−1.48	−1.02	−2.33			

It is known that the MP2 method significantly overestimates the stabilization energy, while the MP3 method underestimates it. To include the higher-order perturbation effects needs the considerable computational costs even by one order. Therefore, we rely on the MP2.5 method [20,21] in the following analysis to represent the interaction energies of amino acid residues, which would provide a quantitative measure.

3.2. Fluctuations in monoclonal antibody recognition

In the analysis of inter-domain interaction energies, we have found that the effect of cross-linked interaction energies should be assessed as well. To gain the information on the molecular recognition by hydrogen bonds in epitopes, we investigate the structures (Fig. 2) and the interaction energy values (Table 2) of the HA antigen–antibody complex. (The signatures (A)–(J) correspond with each other between Fig. 2 and Table 2.)

As seen in Fig. 2(A) and (B) with Table 2(A) and (B), the residue ASP104 in the Fab(I) is bonded by the oxygen and hydrogen atoms to the residue ASN137 in HA(I) with the distance of 1.39 Å, where the calculation result for IFIE is −22.70 kcal/mol. (Hereafter, the distance represents that between hydrogen-bonded donor and acceptor atoms.) On the other hand, the residues ASP104 and SER32 in the Fab(II) are bonded to the residue ASN137 in HA(II) with 1.54 Å and 1.98 Å, respectively; the calculation results for the corresponding IFIEs are −34.54 kcal/mol and −8.70 kcal/mol, respectively. As observed in Fig. 2(C) and (D) with Table 2(C) and (D), the residue ALA61 in Fab(I) interacts attractively with GLY158 and SER159 in HA(I), while both the residues VAL59 and GLY58 in Fab(II) interact with GLY158 and SER159 in HA(II). We also see in Fig. 2(E) and (F) with Table 2(E) and (F) that the interactions between THR192 and SER57 are attractive and repulsive in HA(I)–Fab(I) and HA(II)–Fab(II) complexes, respectively. As an exceptional case in the calculation, as seen in Fig. 2(J) and Table 2(J), LEU164 instead of VAL163 interacts strongly with the Fab by the carbonyl oxygen because of the fragmentation made at C_α, which is employed in the standard FMO recipe [18,19]. In this way, as observed in Fig. 2 and Table 2, the energy values are calculated to be about −5 to −15 kcal/mol with forming the hydrogen bonds shorter than 2 Å, while other electrostatic interaction energy values like the interaction with ASP are larger in magnitude than them. Although the monoclonal antibodies recognize the identical region in HA, the Fab monomers interact with HA trimer with fluctuations. In addition, it is remarked that the present analysis is based on a stable snapshot structure obtained in terms of X-ray crystallographic experiment. Thermal effects at physiological temperatures may thus cause considerable fluctuations in molecular structures and associated interactions of complexes. We should take into account these circumstances in specifying the important amino-acid residues in the HA antigen.

3.3. Hypothetical scheme for predicting the mutations in HA

For the probable mutations of amino acid residues in HA, the following two conditions should be satisfied [10]: that is, the mutant HA should preserve its viral function and also be able to escape the antibody pressure. The former condition is associated with the experimental work carried out by Nakajima et al. [11,12], in which they have extensively introduced single-point mutations in HA and measured the hemadsorption activity of the mutants to assess whether the mutated sites are allowed (positive) or prohibited (negative). The latter condition is associated with the theoretical work in which attractive or repulsive interaction energies with the Fab dimer are evaluated in terms of the values of IFIE sum of the residues in the HA antigenic regions A and B (Table 3). Our hypothesis [10] is that the residues satisfying these two conditions above (i.e., allowed site and attractive interaction) will have a high probability of mutation, which will be examined below through comparison with the historical facts concerning the actual mutations in HA.

Table 3 shows the data regarding the positive(p)/negative(n) sites and the interaction energies with the Fab dimer at the MP2.5/6-31G level for all the amino acid residues in the HA antigenic regions A and B. At first, there are 21 residues of allowed and attractive sites which may be predicted to lead to mutations in our scheme. We see that 17 residues (represented by the symbol *) of them have already been mutated. The other four residues (represented by the symbol **) may be expected to be mutated in future. Next, there are 14 residues at allowed and repulsive sites. Three residues (represented by the symbol ††) of them have not mutated corresponding to our criterion for mutation. However, there are eleven residues (represented by the symbol †) seemingly against our prediction, which are located at the repulsive sites but have experienced the mutations. We will discuss and explain these cases below, where we will categorize them into the charged, polar and hydrophobic residues for quantitative characterization of amino acid residues. In addition, there are a few exceptions concerning the criterion above by the hemadsorption experiment; LYS140 and VAL 202 experienced the mutations in spite of prohibited (n) sites, which may be explained in terms of the differences of mutations between those introduced in the hemadsorption experiment [11,12] and observed actually. Fig. 3 shows the IFIE sums between Fab dimer and each HA monomer in antigenic regions A and B. It is noted that the red bars represent the sites that are allowed (p: positive) and show an attractive interaction

with Fab dimer, which will have a high probability of forthcoming mutation.

As shown in Fig. 3(A), the charged residues LYS156 and ASP172 have possibility to mutate because of positive (allowed) sites and attractive interactions with Fab dimer. In fact, these sites have mutated many times until now. Interaction of LYS156 in HA(III) is very repulsive against Fab dimer because of its non-bonding to Fab dimer with positive charge. However, it shows the large attractive interaction values in HA(I) and (II) so that LYS156 experiences the antibody pressure by the Fab monomers. It is also noted that K156E is favorable for receptor binding and was actually selected under the pressure of antibodies [29]. Interaction energies of charged residues were quantitatively too large due to the neglect of screening effect in the present FMO calculation in vacuo,

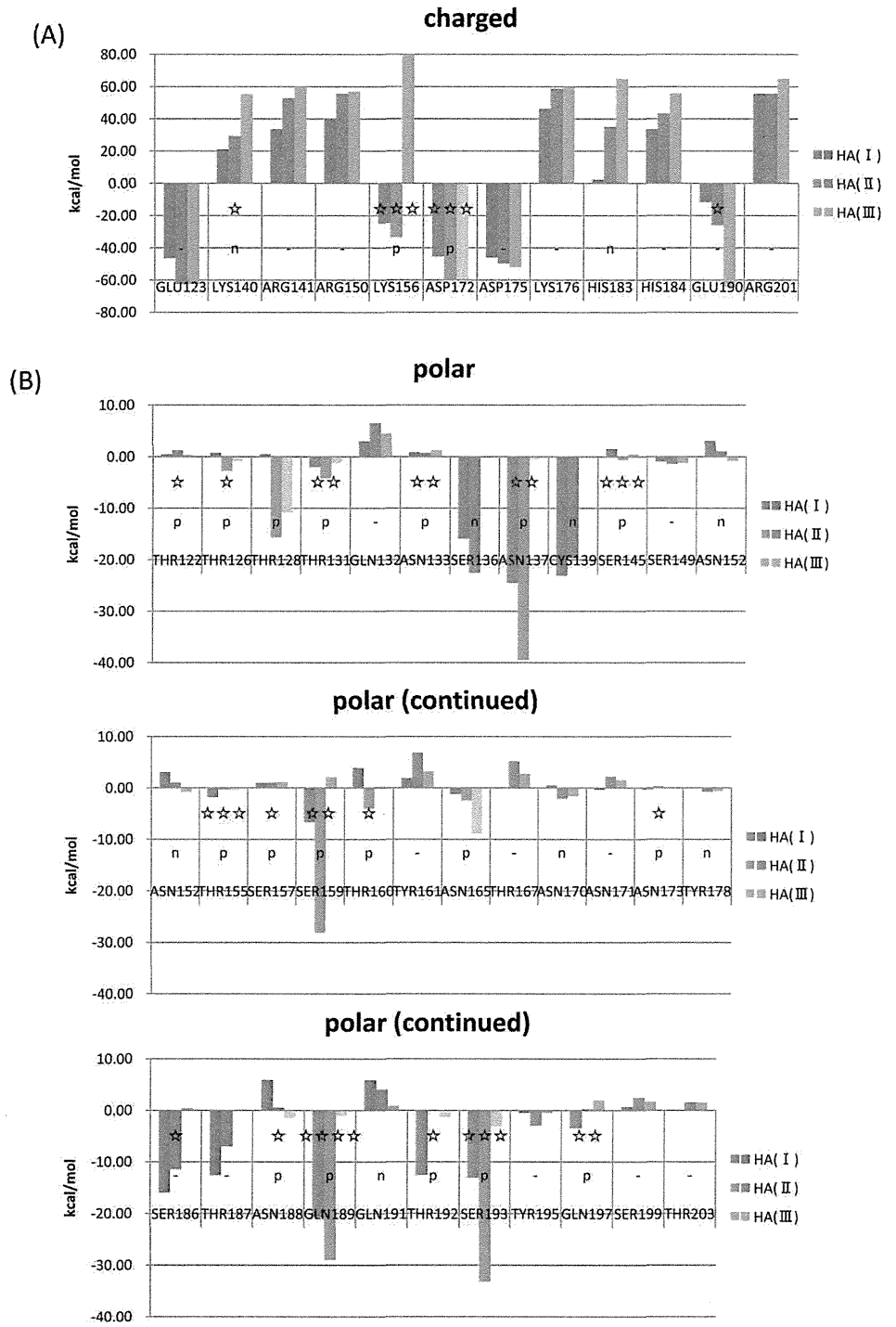

Fig. 3. IFIE sums between Fabs and the residues in the antigenic regions A and B calculated with FMO-MP2.5/6-31G method: (A) charged residues; (B) polar residues; (C) hydrophobic residues. The red bars represent the sites that are allowed (p: positive) and show an attractive interaction with Fab dimer, which will have a high probability of forthcoming mutation; other cases are depicted by the blue bars. The number of stars represents the times of mutations observed already. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

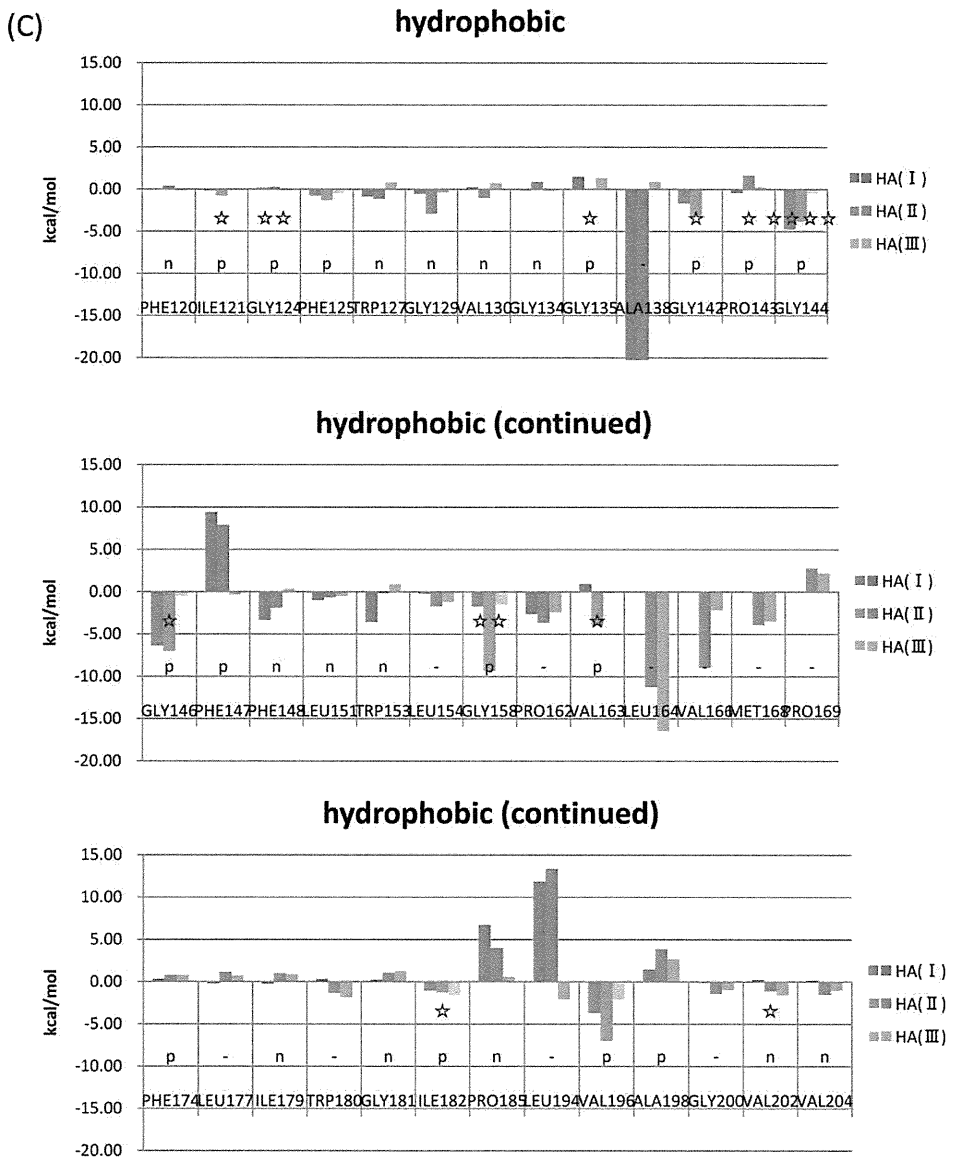

Fig. 3. (*Continued*).

while the comparative features (relative importance and attractive or repulsive interaction) of individual residues are supposed to be described relevantly.

As seen in Fig. 3(B), there are 18 polar residues located at positive (allowed) sites. The eleven sites (126, 128, 131, 137, 155, 159, 165, 173, 189, 192 and 193) are under the antibody pressure apparently, 10 of which have already mutated. The residue 128 may be expected to be mutated in future. Although seven residues (122, 133, 145, 157, 160, 173 and 188) show repulsive interactions with Fab dimer, they have been mutated. The mutations at the 122 and 133 sites have obtained the ability of the oligosaccharide attachment by mutations [30,31]. The residue 133 was actually recognized by monoclonal antibodies until recently [32]. The substitutions N145K and N188D have caused structural changes which allow for the escape from the neutralization by antibodies [33,34]. The substitution S157L allowed for the escape from the antibody HC19 [15]. THR160 receives opposite interactions from multiple Fab monomers. In this case, there are some bonding fluctuations concerning the HA(I) and HA(II) interactions with Fab monomers. We thus suggest that THR160 is under the pressure of an antibody. Because of very small interaction with Fab dimer, the residue ASN173 could not be treated by our prediction scheme.

As observed in Fig. 3(C), the hydrophobic residues show smaller interaction energies than the charged and polar residues. In the present electron-correlated FMO calculations, we can consider the dispersion interactions quantitatively. There are 15 hydrophobic residues located at the allowed (positive) sites. Nine residues (121, 125, 142, 144, 146, 158, 163, 182 and 196) are under the antibody pressure apparently, and seven of them have already been mutated. The residues PHE125 and VAL196 may be expected to be mutated in future. Although three residues (124, 135 and 143) show repulsive interactions with Fab dimer, they have been mutated. The substitution G135R enhances the attractive interaction with glycoprotein [29], and G135T enhances the attractive interaction with sialic acid [35]. The remaining residues GLY124 and PRO143 interact with Fab monomers by 0–1 kcal/mol, which are very weak interactions.

Thus, we have obtained satisfactory results in fair agreement with the historical mutation events, as well as in the earlier study [10]. These results above which were calculated in vacuo, however, ignore the screening effects by solvents [36,37]. Especially,

the IFIE values between charged fragments are substantially overestimated. To cope with this difficulty, although it costs much more, it would be desirable to incorporate the solvation effects into the FMO calculation with, e.g., the Poisson–Boltzmann equation [38]. If this method would be combined with the IFIE calculations, we can evaluate the interaction energies more quantitatively, while careful treatments for the protonation state and the counterions would be required. (Standard aqueous solution pK_a values are assumed in the present analysis.) Further issues to be investigated are the contributions associated with entropy and electronic polarization effects [36,37] in the complex. On the other hand, concerning the calculation levels, it has been observed [28,39] that the relative ordering of the binding energies can be obtained even at such low level as HF/STO-3G. In addition, we can further assess the relative importance of energy values by categorizing the charged, polar and hydrophobic residues, while the interaction energies associated with the hydrophobic residues should be addressed by taking into account the electron correlation effects appropriately.

4. Concluding remarks

The influenza vaccines have been developed mainly against target HA proteins, while it has been difficult to predict forthcoming mutations in HA. We note that only the current calculations cannot detect probable mutations, which may also be associated with the maintenance of HA functions. However, in our combination of computational and experimental methods, to overcome this difficulty, we attempted the detection of possible amino-acid mutations to escape from the antibodies with the aid of the measurement of hemadsorption activities in antigenic regions. As a result, our prediction scheme was found to be consistent with the historical facts of mutations by 83%. (We have picked up the 21 residues of allowed and attractive sites, and two more ones, LYS156 and THR160, showing the attractive interactions at HA monomer level, of which 19 sites have actually mutated.) On the other hand, some mutations unexplained in this analysis were found: however, for example, the substitutions N145K and N188D seemed to induce the structural changes between the HAs of antigenic mutant and the wild type, and G135R and K156E enhanced interactions with glycoprotein under the monoclonal antibodies. Almost all exceptional mutations against our prediction could be explained by considering these previous studies in the literature. Thus the challenge will now be focused on the elucidation of the possible functions of glycoproteins to identify these exceptions. In addition, amino acid changes at more than one residue on the same antigenic site are observed [40,41] for antigenic drift, whose elucidation associated with some cooperative effects may be another challenge. Nonetheless, we expect that this novel computational approach addressed in the present work would be useful for influenza vaccine developments as well.

Acknowledgments

This work was partially supported by the Health and Labour Sciences Research Grants on Emerging and Re-emerging Infectious Diseases (to E.N., No. H22-//Shinko-Ippan-006//) from the Ministry of Health, Labour and Welfare of Japan.

References

[1] J.J. Skehel, D.J. Stevens, R.S. Daniels, A.R. Douglas, M. Knossow, I.A. Wilson, D.C. Wiley, A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody, Proc. Natl. Acad. Sci. U.S.A. 81 (1984) 1779–1783.
[2] R.S. Daniels, A.R. Douglas, J.J. Skehel, D.C. Wiley, Analyses of the antigenicity of influenza haemagglutinin at the pH optimum for virus-mediated membrane fusion, J. Gen. Virol. 64 (1983) 1657–1662.
[3] D.C. Wiley, I.A. Wilson, Structural identification of the antibody-binding site of Hong Kong influenza haemagglutinin and their involvement antigenic variation, Nature 289 (1981) 373–378.
[4] T. Bizebard, B. Gigant, P. Rigolet, B. Rasmussen, O. Diat, S. Bosecke, S. Wharton, J.J. Skehel, M. Knossow, Structure of influenza virus hemagglutinin complexed with a neutralizing antibody, Nature 376 (1995) 92–94.
[5] D. Fleury, B. Barrere, T. Bizebard, R.S. Daniels, J. Skehel, M. Knossow, A complex of influenza hemagglutinin with a neutralizing antibody that binds outside the virus receptor binding site, Nat. Struct. Biol. 6 (1999) 530–534.
[6] D. Fleury, R.S. Daniels, J.J. Skehel, M. Knossow, T. Bizebard, Structural evidence for recognition of a single epitope by two distinct antibodies, Struct. Funct. Gen. 40 (2000) 572–578.
[7] C. Barbey-Martin, B. Gigant, T. Bizebard, L.J. Calder, S.A. Wharton, J.J. Skehel, M. Knossow, An antibody that prevents the hemagglutinin low pH fusogenic transition, Virology 294 (2002) 70–74.
[8] W.D. Kundin, Hong Kong A-2 influenza virus infection among swine during a human epidemic in Taiwan, Nature 228 (1970) 857.
[9] C.A. Russell, The global circulation of seasonal influenza A (H3N2) viruses, Science 320 (2008) 340–346.
[10] K. Takematsu, K. Fukuzawa, K. Omagari, S. Nakajima, K. Nakajima, Y. Mochizuki, T. Nakano, H. Watanabe, S. Tanaka, Possibility of mutation prediction of influenza hemagglutinin by combination of hemadsorption experiment and quantum chemical calculation for antibody binding, J. Phys. Chem. B 113 (2009) 4991–4994.
[11] K. Nakajima, E. Nobusawa, K. Tonegawa, S. Nakajima, Restriction of amino acid change in influenza a virus H3HA: comparison of amino acid changes observed in nature and in vitro, J. Virol. 77 (2003) 10088–10098.
[12] K. Nakajima, E. Nobusawa, A. Nagy, S. Nakajima, Accumulation of amino acid substitutions promotes irreversible structural changes in the hemagglutinin of human influenza AH3 virus during evolution, J. Virol. 79 (2005) 6472–6477.
[13] A.J. Caton, G.G. Brownlee, The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (HI Subtype) , Cell 31 (1982) 417–427.
[14] D.C. Wiley, The structure and function of the hemagglutinin membrane glycoprotein of inflenza virus, Ann. Rev. Biochem. 56 (1987) 365–394.
[15] N.G. Wrigley, E.B. Brown, R.S. Daniels, A.R. Douglas, J.J. Skehel, D.C. Wiley, Electron microscopy of influenza hemagglutinin-monoclonal antibody complexes, Virology 131 (1983) 308–314.
[16] BioStation Viewer, Available at http://www.fsis.iis.u-tokyo.ac.jp/en/result/software/.
[17] R.S. Daniels, S. Jeffries, P. Yates, G.C. Schild, G.N. Rogers, J.C. Paulson, S.A. Wharton, A.R. Douglas, J.J. Skehel, D.C. Wiley, The receptor-binding and membrane-fusion properties of influenza virus variants selected using anti-hemagglutinin monoclonal antibodies, EMBO J. 6 (1987) 1459–1465.
[18] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, M. Uebayasi, Fragment molecular orbital method: an approximate computational method for large molecules, Chem. Phys. Lett. 313 (1999) 701–706.
[19] D.G. Fedorov, K. Kitaura, Extending the power of quantum chemistry to large systems with the fragment molecular orbital method, J. Phys. Chem. A 111 (2007) 6904–6914.
[20] Y. Mochizuki, K. Yamashita, K. Fukuzawa, K. Takematsu, H. Watanabe, N. Taguchi, Y. Okiyama, M. Tsuboi, T. Nakano, S. Tanaka, Large-scale FMO-MP3 calculations on the surface protein of influenza virus, hemagglutinin (HA) and neuraminidase (NA), Chem. Phys. Lett. 493 (2010) 346–352.
[21] M. Pitoňák, P. Neogrády, J. Černý, S. Grimme, P. Hobza, Scaled MP3 non-covalent interaction energies agree closely with accurate CCSD(T) benchmark data, Chem. Phys. Chem. 10 (2009) 282–289.
[22] K. Kitaura, T. Sawai, T. Asada, T. Nakano, M. Uebayasi, Pair interaction molecular orbital method: an approximate computational method for molecular interactions, Chem. Phys. Lett. 312 (1999) 319–324.
[23] K. Fukuzawa, Y. Komeiji, Y. Mochizuki, A. Kato, T. Nakano, S. Tanaka, Intra- and intermolecular interaction between cyclic-AMP receptor protein and DNA: ab initio fragment molecular orbital study, J. Comput. Chem. 27 (2006) 948–960.
[24] M. Ito, K. Fukuzawa, Y. Mochizuki, T. Nakano, S. Tanaka, Ab initio fragment molecular orbital study of molecular interactions between liganded retinoid X receptor and its coactivator: roles of helix 12 in the coactivator binding mechanism , J. Phys. Chem. B 111 (2007) 3525–3533.
[25] I. Kurisaki, K. Fukuzawa, Y. Komeiji, Y. Mochizuki, T. Nakano, J. Imada, A. Chmielewski, S.M. Rothstein, H. Watanabe, S. Tanaka, Visualization analysis of inter-fragment interaction energies of CRP-cAMP-DNA complex based on the fragment molecular orbital method, Biophys. Chem. 130 (2007) 1–9.
[26] T. Iwata, K. Fukuzawa, K. Nakajima, S. Aida-Hyugaji, Y. Mochizuki, H. Watanabe, S. Tanaka, Theoretical analysis of binding specificity of influenza viral hemagglutinin to avian and human receptors based on the fragment molecular orbital method, Comput. Biol. Chem. 32 (2008) 198–211.
[27] Y. Mochizuki, K. Yamashita, T. Murase, T. Nakano, K. Fukuzawa, K. Takematsu, H. Watanabe, S. Tanaka, Large scale FMO-MP2 calculations on a massively parallel-vector computer , Chem. Phys. Lett. 457 (2008) 396–403.
[28] K. Fukuzawa, Y. Mochizuki, S. Tanaka, K. Kitaura, T. Nakano, Molecular interactions between estrogen receptor and its ligand studied by ab initio fragment molecular orbital method , J. Phys. Chem. B 110 (2006) 16102–16110.
[29] P.A. Underwood, J.J. Skehel, D.C. Willy, Receptor-binding characteristic of monoclonal antibody-selected antigenic variants of influenza virus, J. Virol. 61 (1987) 206–208.
[30] M. Knossow, J.J. Skehel, Variation and infectivity neutralization in influenza, Immunology 119 (2006) 1–7.

[31] Y.P. Lin, V. Gregory, M. Bennett, A. Hay, Recent changes among human influenza viruses, Virus Res. 103 (2004) 47–52.
[32] J. Okada, N. Ohshima, R. Kubota-Koketsu, Y. Iba, S. Ota, W. Takase, T. Yoshikawa, T. Ishikawa, Y. Asano, Y. Okuno, Y. Kurosawa, Localization of epitopes recognized by monoclonal antibodies that neutralized the H3N2 influenza viruses in man, J. Gen. Virol. 92 (2011) 326–335.
[33] D.J. Smith, A.S. Lapedes, J.C. de Jong, T.M. Bestebroer, G.F. Rimmelzwaan, A.D.M.E. Osterhaus, R.A.M. Fouchier, Mapping the antigenic and genetic evolution of influenza virus, Science 305 (2004) 371–376.
[34] M. Knossow, R.S. Daniels, A.R. Douglas, J.J. Skehel, D.C. Willey, Three-dimentional structure of an antigenic mutant of the influenza virus haemagglutinin, Nature 311 (1984) 678–680.
[35] A.I. Karasin, M.M. Schutten, L.A. Cooper, C.B. Smith, K. Subbarao, G.A. Anderson, S. Carman, C.W. Olsen, Genetic characterization of H3N2 influenza viruses isolated from pigs in North America, 1977–1999: evidence for wholly human and reassortant virus genotypes, Virus Res. 68 (2000) 71–85.
[36] T. Sawada, D.G. Fedorov, K. Kitaura, Role of the key mutation in the selective binding of avian and human influenza hemagglutinin to sialosides revealed by quantum-mechanical calculations, J. Am. Chem. Soc. 132 (2010) 16862–16872.
[37] T. Sawada, D.G. Fedorov, K. Kitaura, Binding of influenza A virus hemagglutinin to the sialoside receptor is not controlled by the homotropic allosteric effect, J. Phys. Chem. B 114 (2010) 15700–15705.
[38] H. Watanabe, Y. Okiyama, T. Nakano, S. Tanaka, Incorpolation of solvation effects into fragment molecular orbital calculation with the Poisson–Boltzmann equation, Chem. Phys. Lett. 500 (2010) 116–119.
[39] K. Fukuzawa, K. Kitaura, M. Uebayashi, K. Nakata, T. Kaminuma, T. Nakano, Ab initio quantum mechanical study of the binding energies of human estrogen receptor α with its ligands: an application of fragment molecular orbital method, J. Comput. Chem. 26 (2005) 1–10.
[40] I.A. Wilson, N.J. Cox, Structural basis of immune recognition of influenza virus hemagglutinin, Annu. Rev. Immunol. 8 (1990) 737–771.
[41] J.B. Poltkin, J. Dushoff, S.A. Levin, Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 6236–6268.

Sialic Acid Recognition of the Pandemic Influenza 2009 H1N1 Virus: Binding Mechanism Between Human Receptor and Influenza Hemagglutinin

Kaori Fukuzawa[1,*], Katsumi Omagari[2], Katsuhisa Nakajima[2], Eri Nobusawa[3] and Shigenori Tanaka[4]

[1]*Mizuho Information & Research Institute, Inc., 2-3 Kanda Nishiki-cho, Chiyoda-ku, Tokyo 101-8443, Japan;* [2]*Department of Virology, Medical School, Nagoya City University, 1 Kawasumi, Mizuho-chou, Mizuho-ku, Nagoya 467-8601, Japan;* [3]*National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama, Tokyo 208-0011, Japan;* [4]*Graduate School of System Informatics, Department of Computational Science, Kobe University, 3-11 Tsurukabuto, Nada, Kobe 657-8501, Japan*

Abstract: Quantum mechanical fragment molecular orbital calculations have been performed for receptor binding of the hemagglutinin protein of the recently pandemic influenza 2009 H1N1 (2009/HIN1pdm), A/swine/Iowa/1930, and A/Puerto Rico/8/1934 viruses to α2-6 linked sialyloligosaccharides, as analogs of human receptors. The strongest receptor binding affinity was observed for the 2009/H1N1pdm. The inter-fragment interaction energy analysis revealed that the amino acid mutation of 2009/H1N1pdm, Ser145Lys, was a major cause of such strong binding affinity. Strong ionic pair interaction between the sialic acid and Lys145 was observed only in the 2009/H1N1pdm, in addition to the hydrogen bond between the sialic acid and Gln226 observed in all the HAs. Therefore, pandemic 2009/H1N1pdm has been found to recognize the α2-6 receptor much stronger than the 1930-swine and 1934-human.

Keywords: Pandemic influenza 2009 H1N1 virus (2009/H1N1pdm), influenza hemagglutinin (HA), sialic acid recognition, fragment molecular orbital (FMO) method, quantum mechanical calculation, sialo-sugar chain.

1. INTRODUCTION

The emergence of the pandemic influenza 2009 H1N1 (2009/H1N1pdm) viruses has become a world-wide health concern. This virus spread rapidly to countries worldwide, suggesting the ease of human-to-human transmission [1-6]. Three kind of membrane proteins exist in the surface of influenza virus: hemagglutinin (HA), neuraminidase (NA), and M2 proton channel which play key role in the processes of viral penetration, viral elution, and viral replication, respectively [7-10]. A series of significant progresses in studying influenza virus and these membrane proteins, both experimentally and theoretically, have been reported recently [1-35]. The infection of influenza occurs via a binding of HA protein to terminal sialic acids of glycoproteins as the cellular receptors for influenza virus. Two types of linkage, α2-6 and α2-3, between sialic acid and the penultimate galactose residues of carbohydrate side chains found in nature. Human influenza viruses recognize the α2-6 linked sialyloligosaccharides (α2-6 receptors), avian viruses recognize the α2-3 linked sialyloligosaccharides (α2-3 receptors), and swine viruses recognize both. Such selectivity characterizes the binding specificity of host receptors [7-9, 11] Therefore, it is important to quantitatively evaluate binding affinity of HA to α2-6 receptors for the understanding of the infection of 2009/H1N1pdm virus to human.

*Address correspondence to this author at the Mizuho Information & Research Institute, Inc., 2-3 Kanda Nishiki-cho, Chiyoda-ku, Tokyo 101-8443, Japan; Tel: +81-3-5281-5271; Fax: +81-3-5281-5331;
E-mail: kaori.fukuzawa@mizuho-ir.co.jp

Since the outbreak in April 2009, several theoretical discussions about the surface membrane proteins for 2009/H1N1pdm virus have been made based on the three-dimensional structures [12-15]. The amino acid sequences of 2009/H1N1pdm virus were published from the NCBI Influenza virus resources [16]. The present study is focused on a HA for 2009/H1N1pdm virus, and its crystal structure was also published from the Protein Data Bank [17]. Although the crystal structures of HA-receptor complex have not been published, several discussions of the receptor binding specificity have been reported from both the viewpoints of sequence and steric structure [5]. The existence of a positively charged 'lysine fence' from Lys133, Lys145, and Lys222, and additional receptor contact with Asp225 have been discussed as a potential key interactions for receptor binding properties of 2009/H1N1pdm HA. In addition to such structural discussions, molecular simulations enable quantitative discussions about molecular interactions between HA and receptors. Here, we have performed quantum mechanical (QM) calculations to analyze the receptor binding of HA based on the 3D-structure of the HA-receptor complex. Because the 2009/H1N1pdm virus contains an HA gene segment from North American swine lineages, and the crystal structure of the complex between A/swine/Iowa/1930 (1930-swine) HA and α2-6 receptor, Sialyllacto-N-tetraose c (LSTc), have been reported [18], we have modeled 3D-structures of the complex between A/California/04/2009 (one of 2009/H1N1pdm virus) HA (Cal4 HA) and α2-6 receptor from those of the 1930-swine HA. We have also calculated binding specificity of two HAs of H1N1 virus with

0929-8665/11 $58.00+.00

human α2-6 receptor: A/swine/Iowa/1930 and A/Puerto Rico/8/1934 HA, the template swine HA and a prototype human HA (1934-human), respectively.

We applied the fragment molecular orbital (FMO) method [36, 37], which has been developed for efficient and accurate QM calculations for biomolecules, to the detailed analysis of molecular interactions between HA and sialo-sugar chain receptor. The FMO method has already been applied for molecular interactions in some influenza HA systems: host receptor specificity of HA [38-40] and HA antigen-antibody system [41-43]. Based on the FMO method and the inter-fragment interaction energy (IFIE) analysis [38, 41, 44] detailed interactions among amino acid residues can be analyzed, and it is also possible to specify the amino acids governing the receptor-binding of HA. Here, we report the FMO results for binding properties between the human α2-6 type receptor and H1HAs of 1930-swine, 2009/H1N1pdm, and 1934-human. Many studies have indicated that computational approaches, such as structural bioinformatics [45-47], molecular docking [48,49], predicting drug-target interaction [50], pharmacophore modelling [51], and binding mechanism studies [38-40, 44, 49, 52] can timely provide very useful information and insights for drug development and hence are widely welcome by science community. The present study is attempted to use the QM FMO calculations to investigate the binding mechanism of human receptor with influenza HA in hopes to provide useful insights not only for revealing true binding mechanisms [24,25,38-40] but also for drug design [26, 28, 31, 53-60] and vaccine developments against influenza virus.

2. MATERIALS AND METHODS

The initial atomic coordinates of the complexes between H1HA and α2-6 receptors were obtained from the Protein Data Bank (PDB ID: 1RVT and 1RVZ for 1930-swine and 1934-human HAs, respectively) [18]. For the 1930-swine HA-receptor complex, 484 amino acid residues and five units of sialo-sugar chain (Sia1-Gal2-GlcNAc3-Gal4-GlcNAc5) were employed for simulations. On the other hand, 483 amino acid residues and three units of sialo-sugar chain (Sia1-Gal2-GlcNAc3) were employed for the 1934-human HA-receptor complex. The molecular geometry of 2009/H1N1pdm HA complex was modeled from those of 1930-swine HA: the side chains of 1930-swine HA amino acid residues were replaced with those of 2009/H1N1pdm (Cal4) HA at the corresponding positions based on the sequence alignment (Fig. (**1**)). Classical molecular mechanics

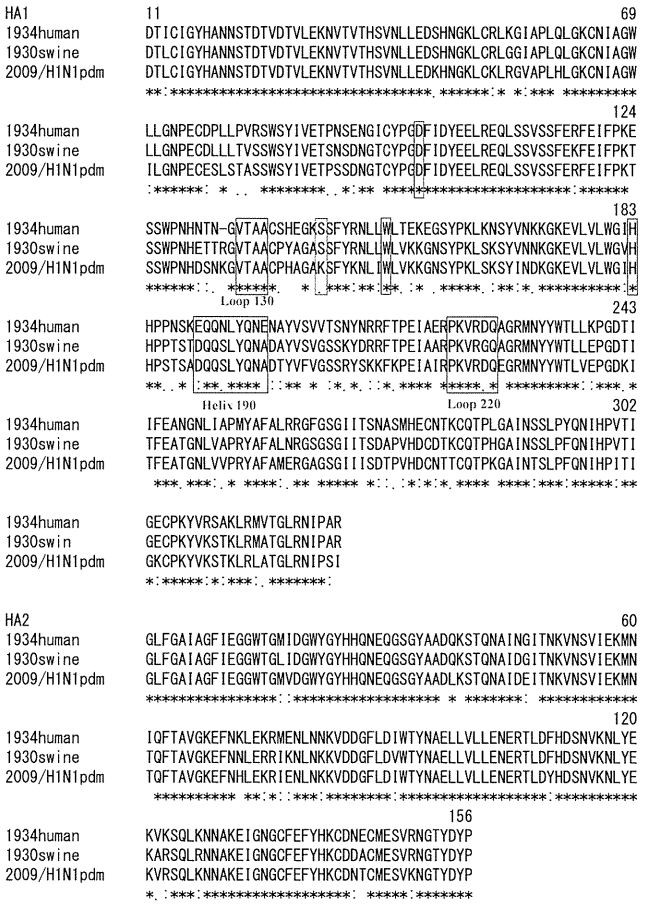

Figure 1. Sequence alignment of 1930-swine, 2009/ H1N1pdm, and 1934-human HAs using the CLUSTALW [69]. '*' indicates fully conserved residue, ':' indicates the residue with strong similarity, and '.' indicates the residue with weaker similarity. Residue numbering is on the basis of H3 HA sequence. The red square indicates 'Ser145Lys' position, The blue squares with rounded corners indicate receptor binding site (Helix 190, Loop130 and Loop220) and conserved residues (Tyr98, Trp153 and His183).