Rational design and engineering of protein A to obtain the controlled elution profile in monoclonal antibody purification

Shinichi Yoshida1*, Dai Murata1, Shunichi Taira1, Keita Iguchi1, Masayuki Takano1, Yoshiyuki Nakano2, Kazunobu Minakuchi2

1Frontier Biochemical & Medical Research Laboratories, Kaneka Corp., 1-8 Miyamae-cho, Takasago-cho, Takasago, Hyogo 676-8688
2QOL Division, Kaneka Corp., 1-8 Miyamae-cho, Takasago-cho, Takasago, Hyogo 676-8688

*E-mail: Shinichi_Yoshida@kn.kaneka.co.jp

(Received July 29, 2011; accepted January 30, 2012; published online February 20, 2012)

Abstract

Biopharmaceutical monoclonal antibodies (Mabs) show different chromatographic behaviors in the elution step on protein A chromatography, although Mabs have similar three-dimensional structures. It is well known that interactions of conventional protein A to the V\textsubscript{H}3 subfamily variable region negatively affect Mabs elution properties. The mutation G29A is known to weaken this binding, although not always sufficiently. We designed novel protein A mutations, S33E and D36R, by a computer-aided evaluation based on the three-dimensional structure. These mutations are expected to not only eliminate protein A binding to the variable region of Mabs but also to maintain its alkaline stability, which is required for effective CIP (Clean in place) of the protein A affinity matrix. In view of the superior potential of C domain, an in vitro study was performed with the G29A mutant of C domain (C-G29A) as a model protein. Both pentameric C domain mutants (C-G29A/S33E.5d and C-G29A/D36R.5d) showed little binding ability to the V\textsubscript{H}3 subfamily variable region of Mabs by BIACORE analysis. We used a C-G29A/S33E.5d-immobilized matrix to confirm that the elution profile of Mabs belonging to the V\textsubscript{H}3 subfamily at pH 3.5 was significantly improved. This matrix also showed almost the same alkaline stability as did the C-G29A.5d-immobilized matrix. The engineered protein A ligand, whose binding ability to the variable region is completely eliminated, would enable the separation of Fab fragments in flow-through fractions from Mab digestions. Rational design by a computer-aided evaluation should enhance the efficiency of protein ligand engineering.

Key Words: molecular mechanics, free energy, protein engineering, protein A, antibody affinity purification

Area of Interest: Molecular Computing
1. Introduction

Protein A affinity chromatography has become the standard process for capturing Mabs from a cell culture supernatant, because it is highly efficient at purifying therapeutic Mabs during industrial production [1][2]. Protein A is a cell-wall associated protein expressed in the Gram-positive bacterium \textit{Staphylococcus aureus}. Protein A consists of a tandem repeat of five highly homologous IgG-Fc binding domains, designated E, D, A, B, and C, and a cell-wall anchoring region, designated XM, from the N terminus [1]. The high binding specificity between protein A and the Fc-region of Mabs enables most of the impurities to be removed in a single step starting directly from a cell culture supernatant.

In protein A affinity chromatography, many parameters affect the chromatographic behavior, such as elution pH, peak shape, and yield. Therefore, it is still difficult to optimize the chromatographic operation for efficient protein separation by this purification process [2][3]. The elution pH of Mabs was reported to be related not only to the structure of the Fc region (subclass designated as IgG1-4), but also to that of the variable region (subfamily designated as V\textsubscript{H}1-8) [3][4]. The binding interface of Mabs to protein A has been localized to the variable region of the Ig heavy chain, but is mainly restricted to V\textsubscript{H}3-encoded human antibodies [5][6][7][8]. The Mab heavy chain V\textsubscript{H}3 subfamily represents nearly half of inherited human V\textsubscript{H} genes and their homologues in other mammalian species [9]. It is well known that this interaction of protein A with the V\textsubscript{H}3-encoded variable region negatively affects the IgG elution properties. In a previous report, non- V\textsubscript{H}3 encoded Mabs had an elution pH of around 3.7-3.8 (estimated with a decreasing pH gradient), while V\textsubscript{H}3-encoded Mabs had a lower elution pH of around 3.1-3.2 [4]. Elution pH is so critical that Mabs exposed to low pH environments risk being damaged chemically and tend to form aggregates, that is, impurities in therapeutic antibodies [2]. These phenomena that are dependent on the kinds of Mabs should be considered when constructing the purification process for the industrial production of therapeutic Mabs.

Z domain is an engineered analog of B domain that was originally developed as an affinity-purification handle for fusion protein production [10]. Z domain contains two amino acid substitutions relative to B domain (A1V and G29A). The G29A mutation especially contributes to improved chemical stability of Z domain by modifying the alkali-susceptible Asn-Gly sequence (at residues 28-29) while adding another feature to B domain. It is reported that Z domain shows a significantly lower affinity for Fabs (polyclonal F(ab)\textsubscript{2} including V\textsubscript{H}3-encoded fragments) as compared with the native IgG-binding domains [11]. MabSelect SuRe, a commercialized protein A-based affinity matrix, comprises a tetramer of modified Z domains [4]. This affinity matrix enables the V\textsubscript{H}3-encoded Mabs to be eluted at a milder pH than the native protein A-based matrix. Thus, the use of the engineered protein A as affinity ligand is thought to be the most effective approach to improve the elution profile of V\textsubscript{H}3-encoded Mabs in protein A chromatography. The G29A mutation is the only well known representative modification that weakens the interaction between protein A and V\textsubscript{H}3-encoded Mabs. The interaction mechanism at the molecular level has been characterized by the X-ray crystal structure analysis of D domain complexed with the Fab fragment of a V\textsubscript{H}3-encoded human IgM antibody [12]. In this complex, the G29A mutation would cause a steric constraint on the interaction, whereby the C\textalpha of Gly-29 is less than 3.5 Å away from the surface of IgM (Gln-H81 and Asn-H82a).

Since Z domain was well characterized as the engineering method to weaken the binding ability of protein A to Fab [4][13], less attention has been paid to other modification methods to regulate this binding ability. For example, C domain, including the G29A mutation, appears to partially retain the binding ability to a V\textsubscript{H}3-encoded variable region. C domain is well known to possess superior resistance to alkaline conditions compared with the other domains, originally, because the
C domain sequence contains a predictable feature to enhance its alkaline stability. There are experimental data showing that the mutation N23T (introduced to Z domain) increases alkaline stability, which supports the high alkaline stability of the C domain having Thr-23 relative to Asn-23 of the other domains [14]. Thus, the engineered protein A ligand based on C domain could provide superior chemical stability for recycle use, but only poorly improve the elution profile of VH3-encoded Mabs.

The present study was conducted to eliminate the binding ability of protein A to a VH3-encoded variable region of Mab by the protein engineering approach. This report describes our attempt to accelerate protein engineering by computer-aided methods. The novel mutations were supported by three-dimensional structural modeling and free energy calculation using this modeling structure. In view of the potential for C domain to act as an IgG affinity ligand, the experiments were performed by using the G29A mutant of C domain (C-G29A) as a model ligand. The effects of these designed mutations are confirmed by BIACORE and chromatography experiments.

2. Materials and Methods

2.1 Three-dimensional structure modeling

The three-dimensional (3D) structure modeling of a C domain complexed with a Fab fragment, including the variable region of Mabs, was constructed by using atomic coordinates determined by the X-ray crystallographic structure of D domain complexed with a Fab fragment of human IgM (PDB ID code: 1DEE) as a template [12]. The 3D structural model of a wild-type C domain was prepared by a homology modeling method based on the well-known alignment of the individual IgG binding domains of protein A [10]. Several insertions and substitutions of amino acid residues were made with the program “Homology Modeling Professional for HyperChem” (Molfunction, http://www.molfunction.com/). Neither insertions nor substitutions were introduced into such secondary structures as $\alpha$-helices. The preferable rotamers (sets of conformations) in the side-chain were also modeled with the same homology modeling program. The 3D structural model of the wild-type C domain in "free"-state (uncomplexed with Fab) was prepared simply by deleting the atomic coordinates of the Fab fragment.

2.2 Computer-aided design of protein

Computer-aided design of C domain mutants of protein A is based on thermostability calculated by a molecular mechanics method [15]. Thermostability is thought to be an important index parameter not only for the free (uncomplexed) state but also for the complexed state with a partner protein. It was assumed in this study that the thermostability of the protein (or the protein complex) could be evaluated by comparing the difference in the free energy change $\Delta G$ between the energy $G_N$ of the properly folded native state and the energy $G_D$ of the unfolded denatured state [eq.1].

$$\Delta G = G_N - G_D \quad \text{[eq.1]}$$

The $\Delta G$ is, in general, a free-energy change in the molecular system in the folding process of the protein. In this method, the $G_N$ and the $G_D$ were calculated based on a rough approximation [eq.2 to 5].

$$G_N = H_N^{\text{CHARMM}} + G_N^{\text{solvation}} \quad \text{[eq.2]}$$

$$G_D = \sum_m H_m^{\text{CHARMM}} + \sum_m G_m^{\text{solvation}} \quad \text{[eq.3]}$$

$$H_m^{\text{CHARMM}} = H_m^{\text{covalent bond}} + H_m^{\text{non-bonded}} \quad \text{[eq.4]}$$

$$H_m^{\text{non-bonded}} = H_m^{\text{vdW1-4}} + H_m^{\text{vdW1-5}} + H_m^{\text{Elec. 1-4}} + H_m^{\text{Elec. 1-5}} \quad \text{[eq.5]}$$
$G_N$ was the sum of the enthalpy term calculated from the CHARMM molecular mechanics potential energy [16] and the solvation energy term based on the atomic occupancy [17]. The enthalpy term is a simple sum of the covalent bond term and the non-bonded term. The parameters of the solvation energy were determined by using experimental free energies of solvation obtained from the octanol-water partition coefficient [18][19]. $G_D$ was approximated from a coordinate of the extended structure of a glycine-rich peptide including the mutated amino acid Xxx, (Gly)$_4$-Xxx-(Gly)$_4$. $G_D$ covers only the mutated amino acid of a mutant, resulting in neglecting the $G_D$ term in $\Delta G$ of the wild type ($\Delta G_{(W)}$).

Here, it is natural to consider the change in the conformational entropy incurred by fixation of the main-chains and side-chains in the folding process. In this study, this change in conformational entropy ($\sum_m T \Delta S_m^{\text{conf}}$) was applied only for the mutated amino acid of a mutant protein. Hence, the $\Delta G$ of a mutant ($\Delta G_{(M)}$) is strictly expressed by the following equation [eq.6].

$$\Delta G_{(M)} = G_N(M) - G_D(M) - \sum_m T \Delta S_m^{\text{conf}}$$ [eq.6]

The residues to mutate were selected by preliminary modeling of the 3D structures of several mutants with the molecular graphic program Swiss PDB Viewer (http://spdbv.vital-it.ch/) [20]. These preliminary modeling efforts were focused on the residues located near the contact interface of C domain with Fab. The subsequent in silico mutant screens were semi-automatically performed with the program Shrike [15]. Following the generation of mutant protein sequences from a wild-type protein sequence, 3D structures of those generated mutants were modeled based on a dead-end elimination (DEE) algorithm [21][22]. The structure of a mutant's side chains was optimized by screening representative side-chain rotamers. The side-chain rotamer that satisfies the following equation [eq.5] is iteratively excluded.

$$E(i_r) - E(i_s) + \sum_j \min[E(i_r, j) - E(i_s, j)] > 0$$ [eq.7]

$E(i_r)$ is the intra-residue energy of rotamer $i$, and $E(i_r, j)$ is the inter-residue energy of rotamers $i$ and $j$. In the case that side-chain rotamers were not narrowed down to one candidate for each residue, the side-chain rotamer having the lowest energy was selected by the calculation of all combinations of side-chain rotamers. Finally, the $\Delta G_{(M)}$ of all mutant proteins and $\Delta G_{(W)}$ of the wild type were calculated with the program Shrike.

To compare the thermostability between a mutant and its wild-type protein, it is necessary to compare the energy difference $\Delta \Delta G$ between the folding energies $\Delta G_{(M)}$ and $\Delta G_{(W)}$ of the mutant and the wild type respectively [eq.8].

$$\Delta \Delta G = \Delta G_{(M)} - \Delta G_{(W)}$$ [eq.8]

2.3 DNA constructions

The coding DNA sequence for a pentameric G29A mutant of the C domain (C-G29A.5d) of *Staphylococcal aureus* protein A was designed by reverse transcription of its amino acid sequence. Each DNA sequence of a monomeric C-G29A was slightly different from the other monomers to prevent homologous recombination. The coding dsDNA was originated from synthetic nucleotides (Takara), and two fragments from the coding dsDNA were subcloned into the vector pSL302. The front fragment was restricted with *PstI* and *HindIII*, and ligated into the pSL302 that had been restricted with the same enzymes. The rear fragment was similarly restricted with *HindIII* and *XbaI*, and ligated into the pSL302 that had been restricted with the same enzymes. The S33E mutation was introduced into each DNA sequence of monomeric C-G29A by site-directed mutagenesis. For example, the front fragment in pSL302, including three sequences of monomeric C-G29A, was mutated by three steps of the double primer PCR method using QuickChange (PROMEGA). The front fragment having the S33E mutation in pSL302 was restricted with *PstI* and *HindIII*. The rear fragment having the same mutation in pSL302 was restricted with *HindIII* and *XbaI*. Both
fragments were simultaneously ligated into pNCMO2 (Takara) that had been restricted with \textit{PstI} and \textit{XbaI}, resulting in an expression plasmid of a pentameric G29A/S33E double-mutant of C domain (C-G29A/S33E.5d). The expression plasmid of a pentameric G29A/D36R double-mutant of C domain (C-G29A/D36R.5d) was also constructed in the same manner, except that oligonucleotide primers were used to introduce the D36R mutation. The DNA sequences were confirmed by a standard DNA sequencing method with ABI 3130xl (Applied Biosystems).

2.4 Protein expression and purification

Protein expression was performed with the \textit{Brevibacillus} secretory expression system. As pNCMO2 is a shuttle vector of \textit{Brevibacillus choshinensis} and \textit{E. Coli}, the expression plasmids constructed with \textit{E. Coli} were transformed into \textit{Brevibacillus choshinensis} cells, strain SP3 (Takara). Fermentation was carried out in Sakaguchi flasks at 30 °C in 3SY medium for 70 hours, following preculture at 30 °C in TM medium for 20 hours. Each protein was purified with the AKTA prime plus system (GE Healthcare). The culture supernatant was adjusted to pH 4.0 and applied to an SP Fast Flow column (GE Healthcare) equilibrated with 50 mM sodium acetate (pH 4.0). The target protein was then eluted by a linear gradient from 0 to 1 M NaCl in 50 mM acetate (pH 4.0). Final purification was performed on a DEAE Fast Flow column (GE Healthcare). The target protein, diluted with pure water, was eluted by a linear gradient from 0 to 0.3 M NaCl in 50 mM Tris-HCl (pH 8.0). Each fractionated protein was then diluted again with pure water, and its purity was confirmed by absorbance measurements at 280 nm with the calculated absorbance coefficients (ProtParam tool of the ExPASy Proteomics Server: http://expasy.org/).

2.5 Preparation of V\textsubscript{H}3-encoded Fab fragment

A V\textsubscript{H}3-encoded Fab fragment was isolated from papain digests of the human monoclonal antibody trastuzumab (Chugai, trade name Herceptin). Trastuzumab was dissolved in 0.1 M sodium acetate (pH 5.5) with 2 mM EDTA and 1 mM cysteine. Papain-immobilized agarose (SIGMA, from papaya latex) was suspended in the protein solution, and this suspension was incubated for 8 hours at 37 °C on a rotator for continuous mixing. The reacted solution separated from the papain-immobilized agarose was applied to a Resource S column (GE Healthcare) equilibrated with 50 mM sodium acetate (pH 4.5). After washing with equilibrating buffer, the Fab fragment was eluted and isolated from the Fc fragment by a linear gradient from 0 to 0.5 M NaCl in 50 mM sodium acetate (pH 4.5). The Fab fragment was finally purified through Gel-filtration chromatography by using Superdex 75 (GE Healthcare) equilibrated with 20 mM sodium phosphate (pH 7.4) and 0.15 M NaCl. Here, too, the purification chromatography was performed with the AKTA prime plus system.

2.6 BIACORE analysis

The interactions of protein with both full-length IgG and the Fab fragment were analyzed with a BIACORE 3000 instrument (GE Healthcare). Human IgG (Baxter, trade name Gammagard) was immobilized by amine coupling onto the carboxylated dextran surface of a CM5 sensor chip (GE Healthcare) by using N-hydroxysuccinimide (NHS) and N-ethyl-N’-(3-dimethylaminopropyl) carbodiimide (EDC) chemistry according to the manufacturer’s protocols. A reference lane was chemically blocked with ethanolamine immediately after activation with NHS/EDC. An IgG solution diluted in 10 mM sodium acetate (pH 4.5) was used in this immobilization process, following the preparation of a 1 mg/mL solution of human IgG in 20 mM sodium phosphate (pH
7.4) and 0.15 M NaCl. Samples of three different concentrations (0.1 – 0.5 μM) of each protein (C-G29A.5d, C-G29A/S33E.5d or C-G29A/D36R.5d) were prepared in Biacore running buffer (20 mM sodium phosphate, 0.15 M NaCl, 0.005% surfactant P20, pH7.4) and injected over the surface of CM5 at a flow rate of 20 μL/min. The surface was regenerated with 50 mM NaOH. The data were analyzed with BIA evaluation software. The sensorgrams of the control lane were subtracted from the sensorgrams of the IgG-immobilized lane. A 1:1 Langmuir model was used to calculate apparent values of the association rate constant $k_a$ (M$^{-1}$ s$^{-1}$), the dissociation rate constant $k_d$ (s$^{-1}$), and the affinity constant $K_A$ (M$^{-1}$, $k_a/k_d$). The Langmuir model is an approximation for the interaction, but is suitable since the sensorgram is evaluated by using the same parameters. The difference in “binding” free energy ($\Delta G_{BIND}$) between the mutant and the wild-type was also calculated from $K_A$ [eq.9].

$$\Delta G_{BIND} = \Delta G_{BIND (M)} - \Delta G_{BIND (W)} = -RT \ln \left( \frac{K_A (M)}{K_A (W)} \right)$$

As for the interaction of protein to the Fab fragment, the measurement and analysis were performed in the same manner as for IgG, except for the concentrations of injected proteins and the regeneration agent. Samples of four different concentrations (4 – 32 μM) of each protein were used in this experiment, and 10 mM NaOH was used as the regeneration agent.

2.7 Evaluation on protein immobilized to affinity matrix

The elution property of the $V_{H3}$-encoded Mab of each protein (C-G29A.5d or C-G29A/S33E.5d) was evaluated as an affinity ligand by immobilization to a matrix for affinity purification. Each protein was covalently coupled to porous cellulose beads (JNC, made to order) by a conventional chemoligation method, by which aldehyde groups are coupled with exposed amines on protein [23]. Glucosamine (amino sugar) was covalently introduced into a formyl group-containing porous cellulose beads by reductive amination with sodium borohydride (Wako Chemicals). Following the generation of oxidatively-cleaved glucosamine (having new formyl groups) by the addition of sodium periodate (Wako Chemicals), a protein was covalently immobilized onto porous cellulose beads through a glucosamine-derivative spacer by reductive amination with sodium borohydride. These affinity columns were connected to the AKTA prime plus system and equilibrated with 20 mM sodium phosphate (pH 7.4) and 0.15 M NaCl. Two milliliters of trastuzumab (1 mg/mL) was applied to these columns, washed with the same equilibrating buffer at a flow rate of 1 mL/min, and then eluted with 35 mM sodium acetate (pH 3.50 or 3.75). Following the elution process, residual trastuzumab was stripped off with 0.5 M acetate (pH 2.5) and 0.1 M sodium sulfite. These sequential affinity chromatography profiles were monitored by UV detection at 280 nm. The alkaline stability of C-G29A/S33E.5d was simply checked by a cycle use test. A Clean-in-place (CIP) cleaning was integrated between each cycle, and the dynamic binding capacity (dBC) [1] of human IgG was checked after cycles 1, 5, 10, 15, and 20. The CIP agent was 0.5 M NaOH, and the contact time for each CIP was 0.5 hour, resulting in a total exposure time of 10 hours.

3. Results and Discussion

3.1 Rational design of the mutations by in silico studies

By the means of computer-aided molecular calculation, amino acid substitutions in the C domain of protein A were evaluated on the basis of thermostability. As described in section 2.2, thermostability was correlated with the difference in the free energy change $\Delta G$ between the energy
of the folded native state and that of the unfolded denatured state. Our criteria for selecting mutants include the difference in free energy changes, which were calculated by the molecular mechanics method based on the 3D structural models of the C domain/Fab complex and C domain in the free-state (uncomplexed with Fab). If the C domain [mutant] Fab complex shows lower thermostability than the C domain [wild type]/Fab complex, the ability to bind to Fab should be weakened. Besides, if the C domain [mutant] in free-state shows higher thermostability than the C domain [wild type], its stability should be positively affected. C domain mutants were evaluated on the basis of both their binding ability and physicochemical stability.

The 3D structural model of the C domain complexed with Fab was based on the X-ray crystallographic structure of a D domain complexed with the Fab fragment of a human IgM (PDB ID code: 1DEE). The amino acid residue positions to mutate were basically selected from among neighbors of the contact interface with Fab. These positions, having a much greater affect on binding to Fab, were narrowed down by a simple visual confirmation using 3D structural modeling. Here, we finally selected Ser-33 and Asp-36 as the positions to mutate. As shown in Figure 1, the mutations at these positions are expected to directly affect binding to Fab. Gly-29, located near the contact interface with Fab, was not selected in this study, one purpose of which was the effective utilization of the chemical stability of the G29A mutation [1][10].

![Figure 1. Schematic representation of C domain/Fab complex](image)

The side chains of important contact residues of C domain (ribbon) to Fab (ribbon and surface) are shown by VDW representation. This figure was generated by VMD [24].

The kinds of amino acids used for substitution were also narrowed down, resulting in the selection of alanine(A), serine (S), threonine (T), asparagine (D), glutamate (E), histidine (H), arginine (R), asparagine (N), and glutamine (Q). The substitutions to these natural amino acids, except for the wild-type (Ser in residue-33 and Asp in residue-36), were introduced in each position. The other natural amino acids were not selected for the following reasons: proline (P) easily breaks the helical structure and Glycine (G) also easily destabilizes the helical structure by its flexibility. Cysteine (C) and lysine (K) are used for immobilization to the matrix, such that it is not preferable to be immobilized near the binding interface to IgG. Valine (V), leucine (L), isoleucine (I), methionine (M), phenylalanine (F), tyrosine (Y), and tryptophan (W) were so unsuitable that substitutions to these hydrophobic amino acids tended to increase nonspecific binding of impurities to a tandem repeat of C domain.

We performed several sets of computational screenings of mutations with the Shrike program. The computational calculation of free-energy differences was based on the molecular mechanics method using the 3D structures. Shrike enables the generation of the 3D structural model of each C
domain [mutant]/Fab complex and the calculation of free-energy differences ($\Delta G_{\text{complex}}^{M}$). The free-energy difference ($\Delta G_{\text{complex}}^{W}$) for the C domain [wild-type]/Fab complex was calculated in the same manner, and then the free-energy differences ($\Delta \Delta G_{\text{complex}}$) between the mutant and wild type were compared. “Binding” free energy differences are generally used to evaluate molecular interaction systems, but we considered that the stability of the binding partner was favorable for evaluations on this protein-protein interaction (In this paper, “binding” free energy is expressed as $\Delta G_{\text{BIND}}$ to discriminate it from free energy in the folding process). Furthermore, additional calculations were performed by varying the parameters on the dielectric constant. Because the calculation on a molecular-interface is especially affected by the dielectric constant, another calculation was carried out by adjusting a relatively high dielectric constant. Free-energy differences ($\Delta \Delta G_{\text{free}}$) between C domain [mutant] and C domain [wild-type] in the free-state were similarly calculated. The results of calculating the thermostability of the C domain/Fab complex and C domain in free-state, which were mutated at residues Ser-33 and Asp-36, are shown in Table 1.

**Table 1. Calculated thermostability as the result of the mutations at residues 33/36 of C domain**

| Mutated a.a. | Ser-33 $\Delta G_{\text{complex}}$ (kcal/mol) | Ser-33 $\Delta G_{\text{free}}$ (kcal/mol) | Asp-36 $\Delta G_{\text{complex}}$ (kcal/mol) | Asp-36 $\Delta G_{\text{free}}$ (kcal/mol) |
|--------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| A (Ala)      | +13.8 (-12.6)                   | -9.3                            | +12.8 (-2.4)                    | -18.9                           |
| D (Asp)      | -1.0 (-13.0)                    | +7.3                            | wild-type                       |                                 |
| E (Glu)      | +9.0 (+0.9)                     | -6.3                            | +8.4 (-23.9)                    | -9.1                            |
| H (His)      | +12.4 (-7.5)                    | -4.7                            | +0.7 (+9.6)                     | -10.0                           |
| N (Asn)      | +11.9 (-17.8)                   | -4.3                            | +26.2 (+11.2)                   | -5.6                            |
| Q (Gln)      | +6.8 (+4.3)                     | +6.8                            | +25.4 (-6.7)                    | -2.9                            |
| R (Arg)      | -10.2 (+13.3)                   | -13.8                           | +9.8 (+6.7)                     | -14.9                           |
| S (Ser)      | wild-type                       |                                 | +11.1 (+1.4)                    | -9.0                            |
| T (Thr)      | +23.9 (-3.3)                    | +0.9                            | +15.4 (-9.9)                    | -5.0                            |

The results of the mutations at residue Ser-33 or Asp-36 of C domain are shown. As for the complex, another result by adjusting the relatively high dielectric constant is shown in parentheses ( ). The negative $\Delta \Delta G$ of the mutants (Fab complex or free-state) indicate better thermostability than the case of the respective wild type. Hence, in this study, the positive $\Delta \Delta G_{\text{complex}}$ is preferable in a complex to eliminate the binding to Fab, while the negative $\Delta \Delta G_{\text{free}}$ is preferable in the free state to maintain stability. These preferable results are shown in **bold-face type**.

The calculated mutations from Ser-33 to Ala, Glu, His, Asn, Gln, and Thr would destabilize the binding to Fab. As for residue Asp-36, all calculated mutations would destabilize the binding to Fab. In a way, these results are acceptable because the substitutions to larger or smaller amino acids in these positions may easily cause a steric constraint on the interface between C domain and Fab. Additional calculations under the relatively high dielectric constant were performed to narrow down the selection of mutations. We empirically know that the mutations showing similar results in
both of the two calculations most closely reflect the actual interaction. As a result of additional calculations, the mutations from Ser-33 to Glu and Gln were selected to destabilize the binding to Fab, with certainty. In the same way, the mutations from Asp-36 to His, Asn, Arg, and Ser were selected. Candidates for the mutations were further narrowed down by evaluating the results of \( \Delta\Delta G^{\text{free}} \). The mutation from Ser-33 to Glu would be better for maintaining the stability of C domain in the free-state than that to Gln. As for the residue Asp-36, all mutations were calculated to have no negative affect on the thermostability of C domain in the free-state. These results might be reasonable, because residue 36 is located at the C-cap position of the \( \alpha \)-helix, and Asp in this position tends to destabilize the \( \alpha \)-helix [25]. We finally ranked the mutations from Asp-36 to His, Asn, Arg, and Ser in order of the value of \( \Delta\Delta G^{\text{free}} \). The mutation from Asp-36 to Arg should be best in view of C domain stability and the destabilization of the binding of C domain to Fab. Consequently, we designed the mutations from Ser-33 to Glu (S33E) and that from Asp-36 to Arg (D36R) to eliminate the binding ability of C domain to V\( \text{H} \text{3} \)-encoded Fab.

### 3.2 Analysis of the mutants by \textit{in vitro} study

For the S33E and D36R mutations designed by the \textit{in silico} study, we evaluated the properties of these mutations by an \textit{in vitro} study. The G29A mutant of C domain (C-G29A) of protein A was selected as the model to introduce those designed mutations. Binding abilities of those designed mutants of C domain (C-G29A/S33E and C-G29A/D36R) were examined by the BIACORE experiment. The BIACORE system is based on a surface plasmon resonance (SPR) method that allows the measurement of biospecific interactions in real time as changes of mass concentrations on a sensor surface. Those designed mutants of C domain, including C-G29A as a reference, were prepared as pentamers by a \textit{Brevibacillus} secretory expression system. Since protein A consists of five IgG-binding domains, a tandem repeat of IgG-binding domains is suitable for capturing IgG as an affinity ligand. We decided to use a tandem repeat construct directly, rather than a single domain, of those designed mutants, because we aimed at improving the efficiency of protein engineering by the \textit{in silico} study.

The abilities those designed mutants to bind to the full-length IgG were first validated. As the affinity ligand to capture IgG, it is required that the binding ability of those designed mutants to the Fc region of IgG should be equivalent to that of native protein A [1]. In this experiment, each mutant was passed by the IgG-immobilized surface. The association rate constant \( k_a \) (M\(^{-1}\) s\(^{-1}\)), the dissociation rate constant \( k_d \) (s\(^{-1}\)), and the affinity constant \( K_A \) (M\(^{-1}\), \( k_a/k_d \)) were calculated for C-G29A.5d, C-G29A/S33E.5d, and C-G29A/D36R.5d. As shown in Table 2, those parameters of C-G29A/S33E.5d and C-G29A/D36R.5d were almost equivalent to that of C-G29A.5d. The affinity constants \( K_A \) of all variants fell inside the range of 1.0 to 5.0 M\(^{-1}\). These data support the conclusion that both of the S33E and D36R mutations have no effect on the binding ability to (the Fc region of) IgG.

### Table 2. Kinetic parameters for the binding of each mutant to IgG

|                  | \( k_{on} \) (×10\(^5\) M\(^{-1}\) s\(^{-1}\)) | \( k_{off} \) (×10\(^{-3}\) s\(^{-1}\)) | \( K_A \) (×10\(^8\) M\(^{-1}\)) | \( \Delta\Delta G_{BIND} \) (kcal/mol) |
|------------------|-----------------------------------------------|----------------------------------------|----------------------------------|--------------------------------------|
| C-G29A.5d        | 4.1                                           | 0.88                                   | 4.7                              | –                                    |
| C-G29A/S33E.5d   | 3.0                                           | 1.3                                    | 2.3                              | +0.42                                |
| C-G29A/D36R.5d   | 4.6                                           | 1.5                                    | 3.0                              | +0.27                                |
Next, the abilities of the designed variants to bind to the V\textsubscript{H}3-encoded Fab fragment were validated by the same BIACORE experiment. The V\textsubscript{H}3-encoded Fab fragment was prepared by the papain digestion of trastuzumab, and each mutant was passed by the Fab-immobilized surface. It was confirmed that the binding ability of C-G29A.5d remains at $10^5$ M\textsuperscript{-1} order of the affinity constant $K_A$. The binding response of C-G29A/S33E.5d (or C-G29A/D36R.5d) became significantly lower than that of C-G29A.5d at the same concentration. The affinity constant $K_A$ to the V\textsubscript{H}3-encoded Fab fragment of C-G29A/S33E.5d (or C-G29A/D36R.5d) was almost less than one-hundredth of those of C-G29A.5d, as shown in Table 3. The association rate constant $k_a$, especially, decreases by introduction of the designed mutations. The S33E (or D36R) mutation seems to eliminate the binding ability of C-G29A to V\textsubscript{H}3-encoded Fab, because the intensity of the affinity constant $K_A$, about $10^3$ M\textsuperscript{-1}, is thought to be a nonspecific binding. Thus, it is clearly demonstrated that the both S33E and D36R mutations have a decreasing effect on the binding ability of C domain to V\textsubscript{H}3-encoded Fab region of IgG.

### Table 3. The kinetic parameters for the binding of each mutant to V\textsubscript{H}3-encoding Fab

|                      | $k_{on}$ ($\times 10^4$ M\textsuperscript{-1}s\textsuperscript{-1}) | $k_{off}$ (s\textsuperscript{-1}) | $K_A$ ($\times 10^5$ M\textsuperscript{-1}) | $\Delta \Delta G_{BIND}$ (kcal/mol) |
|----------------------|-------------------------------------------------|---------------------------------|---------------------------------|----------------------|
| C-G29A.5d           | 5.8                                             | 0.13                            | 4.4                             | –                     |
| C-G29A/S33E.5d      | 0.056                                           | 0.40                            | 0.014                           | +3.4                 |
| C-G29A/D36R.5d      | 0.22                                            | 0.39                            | 0.057                           | +2.6                 |

The performances as an affinity ligand of IgG were confirmed by chromatography experiments using the protein-immobilized matrix. We used C-G29A/S33E.5d by immobilizing it to the porous cellulose matrix by reductive amination, a conventional immobilization method. The elution profiles of the V\textsubscript{H}3-encoded monoclonal IgG trastuzumab were collected through the standard process of IgG affinity chromatography. Elution of IgG was carried out at pH 3.50 or pH 3.75, and regeneration was carried out at pH 2.5. It is difficult to completely elute V\textsubscript{H}3-encoded IgG at pH 3.50 by using conventional protein A affinity sorbents [3][4]. The chromatogram of IgG affinity purification using the C-G29A/S33E.5d-immobilized matrix is also shown in Figure 2. For

![Figure 2](image-url)
comparison, the chromatogram of the C-G29A.5d-immobilized matrix is superimposed in Figure 2. In the case of C-G29A.5d, IgG was completely eluted at pH 3.50, but the shape of the IgG elution peak was a little broad. At pH 3.75, IgG was not completely eluted from the matrix. In the case of C-G29A/S33E.5d, IgG captured by the C-G29A/S33E.5d-immobilized matrix was completely eluted even at pH 3.75. The shape of the IgG elution peak for C-G29A/S33E.5d was sharper than that for C-G29A.5d. The IgG elution properties of C-G29A/S33E.5d as an affinity ligand were shown to be improved by introduction of the S33E mutation.

The alkaline stability of C-G29A/S33E.5d after the purification cycle, including a CIP-step with 0.5 M NaOH, was also examined. Unlike the case for C–G29A, the conventional protein A loses its binding ability to IgG under this strong alkaline condition. The remaining dynamic binding capacity of the C-G29A.5d-immobilized matrix was greater than 95% after a total alkaline exposure time of 5 hrs, as shown in Figure 3. Hence, it is important that the introduced mutations have little influence on the original alkaline stability of the ligand. The C-G29A/S33E.5d-immobilized matrix showed a comparable remaining dynamic binding capacity to the C-G29A.5d-immobilized matrix after alkaline exposure. In designing the mutation, effects on the stability of C domain were considered based on the calculated thermostability of the C domain in free-state. The S33E mutation was calculated to be energetically-favored for the stability of C domain, so that the mutant should keep its high alkaline stability.

![Figure 3. The remaining IgG binding capacity after repeated CIP treatments](image)

### 4. Conclusion

This study demonstrates that rational design by a computer-aided evaluation should contribute to accelerate the engineering of a protein as an affinity ligand. The interactions of conventional protein A to the V_{H3} subfamily variable region negatively affect Mabs elution properties in the affinity chromatography process. Novel mutations were designed by computer aided methods, including 3D structural modeling, and mutational screening based on comparing differences in free energy change. These designed mutations enabled not only the elimination of a protein’s binding to Mabs, but also the maintenance of their alkaline stability, as intended. In this paper, we focused on the improvement of Mabs elution profiles. We are conducting studies to improve alkaline stability by other mutations. These data will be reported in the near future.
The engineered protein A ligand with its Fab-binding ability eliminated could speed-up the development of the downstream process of industrial Mab production [1][2][3][4]. IgG elution profiles depend little on the kinds of Mabs, which is a great advantage in the nearly established platform for the IgG purification process. This study is also an attempt to provide an alternative usage for the protein A affinity column. The protein A affinity media whose ligand has little ability to bind Fab could be used to completely separate the Fc fragment from the Fab fragment [26]. The Fab fragments could be collected as flow-through when a reaction mixture of Mabs with papain, which cleaves the hinge region of Mabs to make Fc and Fab fragments, is applied onto this engineered protein A affinity column. With this column, the Fab fragments can then be purified by only path-through mode chromatography and ion-exchange chromatography, which increases process efficiency by avoiding the necessity for gel-filtration chromatography. We believe that our engineered protein A affinity ligand can be useful for improved industrial scale purification of Mabs and their fragments.

References

[1] Hober, S.; Nord, K.; Linhult, M. Protein A chromatography for antibody purification. *J. Chromatogr. B*. 2007, 848, 40-47.
[2] Shukla, A. A.; Hubbard, B.; Tressel, T.; Guhan, S.; Low, D. Downstream processing of monoclonal antibodies--Application of platform approaches. *J. Chromatogr. B*. 2007, 848, 28-39.
[3] Ishihara, T.; Kadoya, T.; Yoshida, H.; Tamada, T.; Yamamoto, S. Rational methods for predicting human monoclonal antibodies retention in protein A affinity chromatography and cation exchange chromatography: Structure-based chromatography design for monoclonal antibodies. *J. Chromatogr. A*. 2005, 1093, 126-138.
[4] Ghose, S.; Allen, M.; Hubbard, B.; Brooks, C.; Cramer, S. M. Antibody variable region interactions with Protein A: Implications for the development of generic purification processes. *Biotechnol. Bioengin.* 2005, 92, 665-673.
[5] Schroeder, H. W. Jr.; Hillson, J. L.; Perlmutter, R. M. Structure and evolution of mammalian VH families. *Int. Immunol.* 1990, 2, 41-50.
[6] Potter, K. N.; Li, Y.; Pascual, V.; Capra, J. D. Staphylococcal protein A binding to VH3 encoded immunoglobulins. *Int. Rev. Immunol.* 1997, 14, 291-308.
[7] Potter, K. N.; Li, Y.; Capra, J. D. Staphylococcal protein A simultaneously interacts with framework region 1, complementarity-determining region 2, and framework region 3 on human V_{H}3-encoded Igs. *J. Immunol.* 1996, 157, 2982-2988.
[8] Hakoda, M.; Kamatani, N.; Hayashimoto-Kurumada, S.; Silverman, G. J.; Yamanaka, H.; Terai, C.; Kashiwazaki, S. Differential binding avidities of human IgM for staphylococcal protein A derive from specific germ-line VH3 gene usage. *J. Immunol.* 1996, 157, 2976-2981.
[9] Hillson, J. L.; Karr, N. S.; Oppliger, I. R.; Mannik, M.; Sasso, E. H. The structural basis of germline-encoded V_{H}3 immunoglobulin binding to staphylococcal protein A. *J. Exp. Med.* 1993, 178, 331-336.
[10] Nilsson, B.; Moks, T.; Jansson, B.; Abrahmsen. L.; Elmblad. A.; Holmgren. E.; Henrichson, C.; Jones, T. A.; Uhlen, M. A synthetic IgG-binding domain based on staphylococcal protein A. *Protein Eng.* 1987, 1, 107-113.
[11] Jansson, B.; Uhlen, M.; Nygren, P. All individual domains of staphylococcal protein A show Fab binding. *FEBS Immunol. Med. Microbiol.* 1998, 20, 69-78.
[12] Graille, M.; Stura, E. A.; Corper, A. L.; Sutton, B. J.; Taussig, M. J.; Charbonnier, J. B.; Silverman, G. J. Crystal structure of a *Staphylococcus aureus* protein A domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity. *Proc. Natl. Acad. Sci. U S A.* 2000, 97, 5399-5404.

[13] Linhult, M.; Gulich, S.; Graslund, T.; Simon, A.; Karlsson, M.; Sjoberg, A.; Nord, K.; Hober S. Improving the tolerance of a protein analogue to repeated alkaline exposures using a bypass mutagenesis approach. *Proteins.* 2004, 55, 407-416.

[14] Gulich, S.; Uhlen, M.; Hofer, S. Protein engineering of an IgG-binding domain allows milder elution conditions during affinity chromatography. *J. Biotechnol.* 2000, 76, 233-244.

[15] Morikawa, S.; Nakai, T.; Yasohara, Y.; Nanba, H.; Kizaki, N.; Hasegawa, J. Highly Active Mutants of Carbonyl Reductase S1 with Inverted Coenzyme Specificity and Production of Optically Active Alcohols. *Biosci. Biotechnol. Biochem.* 2005, 69, 544-552.

[16] MacKerell, A. D.; Bashford, D.; Bellott, R. L.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B.* 1998, 102, 3586-3616.

[17] Stouten, P. W.; Froemmel, C.; Nakamura, H.; Sander, C. An Effective Solvation Term Based on Atomic Occupancies for Use in Protein Simulations. *Mol. Simul.* 1993, 10, 97-120.

[18] Fauchère, J.-L.; Pliska, V. Hydrophobic parameters of amino-acid side chains from the partitioning of N-acetyl-amino-acid amides. *Eur. J. Med. Chem.- Chim. Ther.* 1983, 18, 369-375.

[19] Eisenberg, D.; McLachlan, A. D. Solvation energy in protein folding and binding. *Nature.* 1986, 319, 199-203.

[20] Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis.* 1997, 18, 2714-2723.

[21] Desmet, J.; Hazes, B.; Lasters, I. The dead-end elimination theorem and its use in protein side-chain positioning. *Nature.* 1992, 356, 539-542.

[22] Lasters, I.; Maeyer, M. D.; Desmet, J. Enhanced dead-end elimination in the search for the global minimum energy conformation of a collection of protein side chains. *Prot. Engin.* 1995, 8, 815-822.

[23] Wong, L. S.; Khan, F.; Micklefield, J. Selective Covalent Protein Immobilization: Strategies and Applications. *Chem. Rev.* 2009, 109, 4025-4053.

[24] Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Molec. Graphics* 1996, 14, 33-38.

[25] Serrano, L.; Sancho, J.; Hirshberg, M.; Fersht, A. R. Alpha-Helix stability in proteins: I. Empirical correlations concerning substitution of side-chains at the N and C-caps and the replacement of alanine by glycine or serine at solvent-exposed surfaces. *J. Mol. Biol.* 1992, 227, 544-559.

[26] Seldon, T. A.; Hughes, K. E.; Munster, D. J.; Chin, D. Y.; Jones, M. L. Improved Protein-A separation of V(H)3 Fab from Fc after Papain Digestion of Antibodies. *J. Biomol. Tech.* 2011, 22, 50-52.