Engineering streptococcal protein G for increased alkaline stability

Susanne Gülich[1], Martin Linhult[1], Stefan Ståhl[2] and Sophia Hober[1,3]

[1]Department of Biotechnology, Royal Institute of Technology (KTH), Stockholm Centre for Physics, Astronomy and Biotechnology (SCFAB), SE-106 91 Stockholm and [2]Affibody AB, Box 20137, SE-161 02 Bromma, Sweden

[3]To whom correspondence should be addressed.
E-mail: sophia@biotech.kth.se

Most protein-based affinity chromatography media are very sensitive towards alkaline treatment, which is a preferred method for regeneration and removal of contaminants from the purification devices in industrial applications. In a previous study, we concluded that a simple and straightforward strategy consisting of replacing asparagine residues could improve the stability towards alkaline conditions. In this study, we have shown the potential of this rationale by stabilizing an IgG-binding domain of streptococcal protein G, i.e. the C2 domain. In order to analyze the contribution of the different amino acids to the alkaline sensitivity of the domain we used a single point mutation strategy. Amino acids known to be susceptible towards high pH, asparagine and glutamine, were substituted for less-alkali-susceptible residues. In addition, aspartic acid residues were mutated to evaluate if the stability could be further increased. The stability of the different C2 variants was subsequently analyzed by exposing them to NaOH. The obtained results reveal that the most sensitive amino acid towards alkaline conditions in the structure of C2 is Asn36. The double mutant, $C2_{N7,36A}$, was found to be the most stable mutant constructed. In addition to the increased alkaline stability and also very important for potential use as an affinity ligand, this mutated variant also retains the secondary structure, as well as the affinity to the Fc fragment of IgG.
Keywords: affinity chromatography/deamidation/protein G/ purification/stabilization

Introduction

Most protein-based affinity chromatography media show significant fragility towards alkaline conditions. This is a major problem in many industrial applications where it is of paramount importance to be able to remove contaminants from the chromatography media, usually accomplished by integration of a cleaning-in-place (CIP) protocol. In this protocol, sodium hydroxide (NaOH) in concentrations ranging from 0.1 to 1 M is the most commonly used agent, resulting in an extremely harsh environment with high pH (Asplund *et al.*, 2000). The sensitivity of proteins towards such conditions is a drawback when using them as binding ligands, and therefore significant efforts have been put into the development of different synthetic ligands, such as various organic mimics (Fassina *et al.*, 1996; Li *et al.*, 1998). This approach often results in binders with increased alkaline resistance. However, an encountered

drawback is often the significantly decreased affinity of these ligands compared to ligands consisting of protein domains. Also, the possibility to use alternative chromatographic techniques such as ion-exchange chromatography has been explored in many applications (Ford *et al.*, 1991; Gräslund *et al.*, 2000). Despite the many positive characteristics of these powerful techniques, protein-based affinity chromatography would for many applications be the preferred choice due to the high selectivity and affinity that can be obtained, especially if more stable protein ligands were available. Nevertheless, the alkaline sensitivity has been a major drawback and a great challenge to encounter. We have earlier shown that it is possible to stabilize a protein ligand towards alkaline conditions using a protein engineering strategy. This strategy involves replacement of asparagine residues with other amino acids less susceptible towards alkaline conditions (Gülich *et al.*, 2000b).

Asparagine residues have been recognized as the major contributors to the alkaline sensitivity, but also glutamine and to a lesser extent aspartate residues can be chemically modified (Stephenson and Clarke, 1989; Wright, 1991; Tomizawa *et al.*, 1995). These modifications may change the function or the potency of a protein or a peptide. The reaction mechanism involves the main chain peptide nitrogen of the succeeding residue, which functions as the nucleophile and attacks the side chain carbonyl of the asparagine, glutamine or aspartate. This cyclization reaction results in a cyclic imide intermediate, which is hydrolyzed to isomerization products (aspartate/glutamate and isoaspartate/isoglutamate). The resulting isomers may be found in their L- or D-form. In addition, cleavage of the peptide bond can occur. These reactions are spontaneous but the rate is severely increased at alkaline pH. Also, temperature and ionic strength are parameters that influence the rate (Geiger and Clarke, 1987). Different proteins are modified to a different extent since the deamidation/isomerization rate is highly sequence and conformation dependent (Kossiakoff, 1988; Lura and Schirch, 1988; Wearne and Creighton, 1989).

In this paper, we evaluate if the protein engineering strategy used to stabilize the albumin-binding domain (ABD) derived from streptococcal protein G (SPG) (Gülich *et al.*, 2000b) can be used to increase the stability towards alkaline treatment for yet another protein ligand, the C2 domain. This is a well known and thoroughly characterized protein, exploited in commercially available protein-based affinity chromatography media. The C2 domain is derived from SPG, a cell-surface protein from *Streptococcus* G148 (Olsson *et al.*, 1987), and has the potential to be used in large-scale purification of monoclonal antibodies. Streptococcal protein G is a multidomain protein with separated albumin-binding and immunoglobulin-binding regions (Nygren *et al.*, 1988). SPG exhibits a broad spectrum of binding to IgG subclasses, and binds to all four human subclasses of IgG, as well as several animal IgG subclasses (Björck and Kronvall, 1984; Åkerström *et al.*, 1985; Olsson *et al.*, 1987). The IgG-binding region
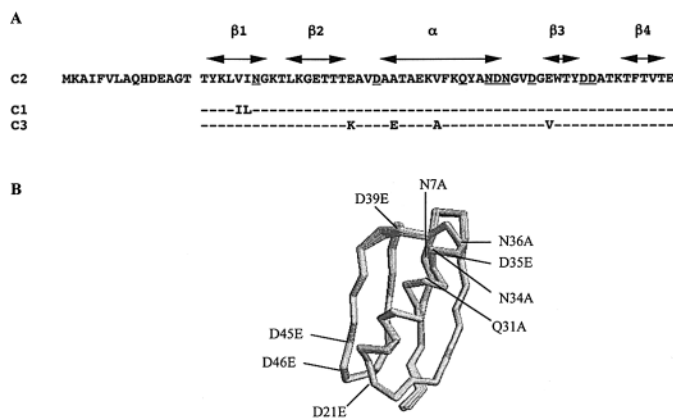
A

B

Fig. 1. (**A**) The 55-residue sequence of the C2 domain (Olsson *et al.*, 1987). Also, the additional 13-residue (asparagines excluded) leader and linker sequence derived from the *Trp*LE expression vector is visualized. The different secondary structures are marked and the substituted residues are underlined. In addition, the sequences for the homologous C1 and C3 domains are shown (Olsson *et al.*, 1987). (**B**) The three-dimensional structure of the C1 domain (Kuszewski *et al.*, 1999). Rasmol was used to generate the picture (Sayle and Milner-White, 1995). The different substitutions are indicated.

consists of three domains, C1, C2 and C3, exhibiting high sequence homology with only a few residue exceptions (Figure 1A). The C2 domain is a small protein constituted of only 55 residues. We have chosen to produce it with an additional leader and linker sequence of 13 amino acids derived from the leader region in the *Trp*LE expression vector (asparagines excluded) (Figure 1A). The structure of C2 and the homologous C1 and C3 domains has been determined both alone and in complex with Fab and Fc by NMR (Gronenborn *et al.*, 1991; Lian *et al.*, 1992) and X-ray crystallography (Achari *et al.*, 1992; Derrick and Wigley, 1992, 1994; Gallagher *et al.*, 1994; Sauer-Eriksson *et al.*, 1995). It is constituted of a four-stranded β-sheet crossed diagonally by a single α-helix. The central two strands comprising the N- and C-termini are parallel. The outer two strands are anti-parallel to the central two strands and connected by the helix (Figure 1B). The C2 domain has no additional stabilizing features such as disulfide bridges or bound prosthetic groups. The topology, the extensive H-bonding network, and the tightly packed hydrophobic core are probably responsible for the high thermal stability. Almost all residues are involved in regular secondary structure elements, which ensure a large number of stabilizing H-bonds (Gronenborn *et al.*, 1991). The domain shows strong interaction to the Fc fragment (Sauer-Eriksson *et al.*, 1995) and also weak interaction to the Fab fragment (Derrick and Wigley, 1992, 1994) of IgG from various species including human. Hence, these characteristics have given the IgG-binding region of SPG a biochemical utility in purifying antibodies, as well as in folding and stability studies.

Goward and coworkers have reported on the sensitivity of protein G to alkaline pH (Goward *et al.*, 1991). In this study, a truncated variant of the IgG-binding region of protein G was incubated at different pH. After incubation at pH 12 for 6 days the IgG-binding capacity was not restored, and circular dichroism (CD) analysis indicated a significant loss of α-helical structure, whereas the β-sheet structure was restored (Goward *et al.*, 1991). Here, we have addressed the fragility of the C2 domain by designing single mutants to resolve the contribution of each asparagine residue to the deamidation

rate. In addition, the possibility of further stabilization by aspartate and glutamine replacements has been investigated.

Materials and methods

DNA constructions and bacterial strains

The plasmid pC2C3 (Nygren *et al.*, 1988) was used as template in a site-directed mutagenesis using a two-step PCR technique (Higuchi *et al.*, 1988). Oligonucleotides coding for the 13-residue leader and linker sequence of the *Trp*LE expression vector (asparagines excluded), the nine-residue linker sequence in the dimers, and the different residue substitutions were synthesized by Interactiva (Interactiva Biotechnologie GmbH, Ulm, Germany). The cloning was performed essentially as described by Sambrook *et al.* (Sambrook *et al.*, 1989), using the restriction enzymes *Xba*I and *Hin*dIII (MBI Fermentas Inc., Amhurst, NY), and the vector pDHZ (Jansson *et al.*, 1996). The sequence of inserted fragments was verified on a MegaBACE™ 1000 DNA Sequencing System, using cycle sequencing based on the dideoxy method (Sanger *et al.*, 1977) with MegaBACE™ terminator chemistry according to the supplier's recommendations (Amersham Biosciences, Uppsala, Sweden). *Escherichia coli* strain RR1ΔM15 (American Type Culture Collection, Rockville, MA) was used during cloning procedures and 017 (Olsson and Isaksson, 1979) was used for expression of the different gene products.

Production and purification

The wild-type protein and mutants thereof were produced according to the protocol described by Gülich *et al.* (Gülich *et al.*, 2000a). After 20 h of cultivation the cells were harvested by centrifugation (10 min at 5000 *g*). The cells were resuspended in 25 ml of cold TST (25 mM Tris–HCl pH 7.5, 150 mM NaCl, 1.25 mM EDTA, 0.05% Tween 20) followed by sonication by Vibra™ cell (Sonics and Materials Inc., Danbury, CT). The soluble gene products were purified as described by Gülich *et al.* (Gülich *et al.*, 2000a). The bound material was eluted with 0.5 M HAc pH 2.9. The amount of protein was estimated by absorbance measurements at 280 nm using the specific absorbance coefficient, a (1 g^{-1} cm^{-1}) (Gill and von Hippel, 1989), (C2 1.244; C2$_{N7A}$, C2$_{N34A}$, C2$_{N36A}$ 1.251; C2$_{N7,36A}$ 1.258; C2$_{N7,36AD35E}$, C2$_{N7,36AD39E}$ 1.255; C2$_{N7,36AD21,45,46E}$ 1.251; C2$_{N7,36AQ31A}$ 1.267; C2dim 1.385; C2$_{N7,36A}$dim 1.402; C2cys 1.227; C2dimcys 1.376; C2$_{N7,36A}$-dimcys 1.392). The concentration was confirmed by analysis of amino acid composition in a Beckman 6300 amino acid analyzer (Beckman, Fullerton, CA) after hydrolysis in 6 M HCl at 155°C for 45 min.

The homogeneity and the molecular weights were confirmed by mass spectrometry using a mass spectrometer (Micromass, Manchester, UK) equipped with an electrospray unit. Calculated and detected molecular weights of the main peak: C2 7663.4 (7663); C2$_{N7A}$, C2$_{N34A}$, C2$_{N36A}$ 7620.4 (7620); C2$_{N7,36A}$ 7577.4 (7577); C2$_{N7,36AD35E}$, C2$_{N7,36AD39E}$ 7591.4 (7591); C2$_{N7,36AD21,45,46E}$ 7619.4 (7619); C2$_{N7,36AQ31A}$ 7520.3 (7520); C2dim 14681.1 (14681); C2$_{N7,36A}$dim 14509.0 (14509); C2cys 7766.6; C2dimcys 14784.3; C2$_{N7,36A}$dimcys 14612.2.

Biospecific interaction analysis

A Biacore™ 2000 instrument (Biacore, Uppsala, Sweden) was used to analyze the affinity of the different variants to the Fc fragment of IgG. Human polyclonal Fc (Jackson ImmunoResearch Laboratories Inc., West Grove, PA) was immobilized on the carboxylated dextran layer of a CM5 sensor chip (research

grade) (Biacore) by utilizing the *N*-hydroxysuccinimide and *N*-ethyl-*N*′-(3-diethylaminopropyl)-carbodiimide chemistry. Immobilization of Fc resulted in ~1500 RU. Six different concentrations (0.1–0.6 µM) of C2 and mutants thereof were prepared in HBS (10 mM HEPES, 0.15 M NaCl, 3.4 mM EDTA, 0.005% surfactant P20, pH 7.4) and injected as duplicates in random order at a flow rate of 30 µl/min. Five millimolar HCl was used to regenerate the surface. The BIA evaluation 3.0.2b software (Biacore) was used to evaluate the data employing a 1:1 (Langmuir) model. This model is an approximation for the interaction but since all variants are evaluated using the same parameters, the kinetic constants are comparable. The signals from a blank surface were subtracted from the Fc surface and apparent k_a (M^{-1} s^{-1}), k_d (s^{-1}) and K_A (M^{-1}) were calculated.

The free thiol group of the C-terminal cysteine in C2cys, C2dimcys and C2$_{N7,36A}$dimcys was used for directed immobilization to a CM5 sensor chip (Biacore) pre-activated with NHS/EDC chemistry according to the supplier's recommendations. 2-(2pyridinyldithio)ethaneamine hydrochloride was injected over the surfaces prior to immobilization of the ligands, which were reduced in 10 mM dithiothreitol followed by desalting on a Sephadex™ G-25 column (Amersham Biosciences) into 30 mM NH$_4$Ac pH 4.0. The immobilization resulted in ~900 RU of each ligand. Human polyclonal Fc and Fab fragments (0.1 and 0.3 µM) (Jackson ImmunoResearch Laboratories Inc.) in HBS buffer were injected as duplicates in random order at a flow rate of 30 µl/min. Ten millimolar HCl was used for regeneration. The signals from a blank surface were subtracted from the ligand surfaces.

Affinity chromatography with a CIP protocol

C2 and mutants thereof were covalently coupled to HiTrap™ columns (Amersham Biosciences) by using the *N*-hydroxysuccinimide chemistry according to the manufacturer's recommendations. An ordinary affinity chromatography scheme was followed on the ÄKTA™ Explorer 10 (Amersham Biosciences). Human polyclonal IgG in TST was loaded on the columns. The columns were washed with TST and 5 mM NH$_4$Ac pH 5.5 before elution with 0.5 M HAc pH 2.8. The eluted material was detected at 280 nm. NaOH (0.1 M) was used as cleaning agent between the different cycles. Each pulse with NaOH was 2 min.

Circular dichroism spectroscopy

A J-720 spectropolarimeter (JASCO, Tokyo, Japan) was used to analyze the secondary structure content. Samples of 0.1 mg/ml were prepared in 10 mM K$_3$PO$_4$ pH 6.0. Spectra were recorded in the far UV region from 250 to 195 nm at 20°C in a quartz cell of pathlength 0.1 cm, and with a scan speed of 20 nm/min. Each spectrum was the mean of three accumulated scans and the final spectra were converted into mean residue ellipticity (deg cm^2 dmol^{-1}).

Thermal denaturation of the different variants was monitored by following the CD signal at 217 nm as a function of temperature. Samples of 0.05 mg/ml were prepared in 10 mM NaAc pH 5.3. This pH was chosen since previous studies of the homologous C1 and C3 domains suggest that the protein is most stable at this pH. However, native structure is maintained at 25°C over a wide pH range (pH 1.5–11.0) (Alexander *et al.*, 1992a,b). The CD signals were measured in a 1-cm quartz cell. The interval between measure points was 5°C. Signals were transformed to apparent fraction of unfolded protein, $F_{app} = (y_f - y_{obs}) / (y_f - y_u)$, where y_f is the value of

the folded state, y_{obs} is the observed value at different temperatures and y_u is the value of the unfolded state, and plotted versus the temperature. $T_m^{1/2}$ was read from the graph at $F_{app} = 0.5$. The temperature was also lowered in order to investigate reversibility of the denaturation.

Chemical denaturation was also monitored by following the signal at 217 nm as a function of the concentration of denaturant. Samples of 0.25 mg/ml were prepared in 10 mM NaAc pH 5.3. A stock solution of 8 M GdnHCl (Sigma Chemical Co., St. Louis, MO) was mixed with the protein solution to achieve different concentrations of denaturant. Measurements took place in a 0.1-cm quartz cell. Signals were transformed to F_{app} values and plotted versus the concentration of denaturant. $C_m^{1/2}$ was read from the graph at $F_{app} = 0.5$.

RP-HPLC analysis after NaOH incubation

Proteins incubated in NaOH were also analyzed by RP-HPLC. After different incubation times in 0.1 M NaOH at room temperature (RT) the samples were injected directly onto a Source™ 5RPC ST4.6/150 column (Amersham Biosciences) suited for high pH. The apparatus was a Hewlett Packard Series II 1090 Liquid Chromatograph (Wilmington, DE). A linear gradient from 30 to 50% of acetonitrile with 0.25% pentafluoropropionic acid for 20 min was used. The peaks were monitored at 280 nm. The eluted peaks were collected and analyzed on a VG Platform mass spectrometer (Micromass) equipped with an electrospray unit.

Isoelectric focusing gel electrophoresis (IEF)

The sensitivity towards deamidation for C2 and C2$_{N7,36A}$ was also investigated using IEF under non-denaturing conditions. The proteins were incubated in 0.1 M NaOH for 2 h at RT. The buffer was changed by gel filtration using Sephadex™ G-25 columns (Amersham Biosciences) equilibrated with milliQ-water. The IEF gel was casted and run on a Model 111 Mini IEF cell (Bio-Rad Laboratories, Hercules, CA) according to the manufacturer's recommendations using Ampholine™ preblended pH 4.0–6.5 (Amersham Biosciences). The IEF calibration kit was purchased from Amersham Biosciences.

Results

In this study, a strategy was designed to investigate the importance of different alkali sensitive residues in the decrease of activity of the C2 domain of SPG when exposed to NaOH. A first generation of mutated variants was assigned to asparagine replacements since this residue is known to be the most sensitive amino acid towards alkaline conditions. The C2 domain contains three asparagine residues, which were substituted for alanines since no alternative amino acids were found in the homologous sequences (Figure 1). The asparagines are preserved in the homologous C1 and C3 domains (Figure 1A). Since alanine is a small amino acid and is not known to introduce any unusual conformational properties in the structure it is suitable for this purpose. Single mutants were designed in order to resolve the contribution of each asparagine to the deactivation rate. These single mutants were designated C2$_{N7A}$, C2$_{N34A}$ and C2$_{N36A}$. Also, one variant with two asparagines exchanged for alanine, C2$_{N7,36A}$, as well as one with three asparagines exchanged for alanine, C2$_{N7,34,36A}$, were constructed. A second generation of mutants was generated using the most stable mutated variant, C2$_{N7,36A}$, as scaffold. Two single mutants and one triple mutant were designed in which the aspartate residues were substituted for glutamate residues,

$C2_{N7,36AD35E,}$ $C2_{N7,36AD39E}$ and $C2_{N7,36AD21,45,46E}$. In order to explore the sensitivity of the single glutamine in the structure, $C2_{N7,36A}$ was used as scaffold again. The glutamine was substituted for alanine resulting in $C2_{N7,36AQ31A}$.

Biospecific interaction analysis of the asparagine mutants

The surface plasmon resonance (SPR) technology was used to investigate the interaction between the Fc fragment of IgG and the C2 variants. This technology gives information on affinities as well as the kinetics of both association and dissociation in real time. Hence, it is well suited for detailed screening of the interactions of modified proteins. The sensorgrams reveal a similar binding behavior for C2, $C2_{N7A}$ and $C2_{N36A}$ (Figure 2A). C2 and $C2_{N7A}$ show very similar kinetic parameters (Table I). This can be expected since Asn7 is situated on the other side of the domain facing away from the Fc-interacting region. The $C2_{N36A}$ variant has a slightly decreased dissociation rate constant indicating an increased stability of the complex with Fc (Table I). The $C2_{N34A}$ mutant shows a very weak interaction with the Fc fragment and no reliable kinetic parameters could be calculated (Figure 2A). Based on these data a double mutant was designed, $C2_{N7,36A}$, preserving Asn34 in the sequence. The kinetic parameters of $C2_{N7,36A}$, were shown to be very similar to that of $C2_{N36A}$. Also, a triple mutant with all three asparagines substituted for alanines was designed for characterization. Interestingly, we were unable to produce this triple mutant in *E.coli* cells. Therefore, a gene fusion strategy was employed, taking advantage of the positive characteristics of the ABD (Kraulis *et al.*, 1996). The idea was to render an affinity purification procedure (Ståhl *et al.*, 1999) possible by fusing the ABD domain C-terminally to $C2_{N7,34,36A}$. However, this strategy did not result in any product, thus indicating that substitution of all asparagines results in a very labile, probably incorrectly folded protein.

Finally, genetically fused dimers of the wild-type C2 and also the double mutant $C2_{N7,36A}$ were designed and expressed in order to investigate the effects on the affinity, when linking two active domains to each other. The linker connecting the monomers was designed using the linker regions connecting the three different C domains of SPG (KPEVYPAVT) (Olsson *et al.*, 1987). The dimers were designated C2dimcys and $C2_{N7,36A}$dimcys. By using directed immobilization the dimers were coupled to the chip surface and Fc was injected over the surface. Also inversely, the dimers were injected over a sensor chip with immobilized Fc. For comparison, the monomeric C2 domain was also equipped with a C-terminal cysteine (C2cys) and coupled to the chip surface. A slower dissociation rate was detected for both dimers as well as the wild-type monomer coupled to the surface compared to the Fc-coupling (Figure 2A and B). Noteworthy, both dimers show significantly higher binding capacity compared to the monomer, C2cys, when the same amount of functional domains is immobilized for monomer and dimers (Figure 2B). Hence, the length of the linker seems sufficient to separate the two monomers and still retain their function. Finally, human polyclonal Fab fragment was injected over the surfaces. According to the sensorgram the C2 domain exhibits very low affinity to Fab compared to Fc (Figure 2C). Interestingly, a loss of the Fab interaction is observed for $C2_{N7,36A}$dimcys (Figure 2C). This is most likely due to the replacement of Asn36, which is suggested to take part in the Fab interaction (Derrick and Wigley, 1992, 1994).
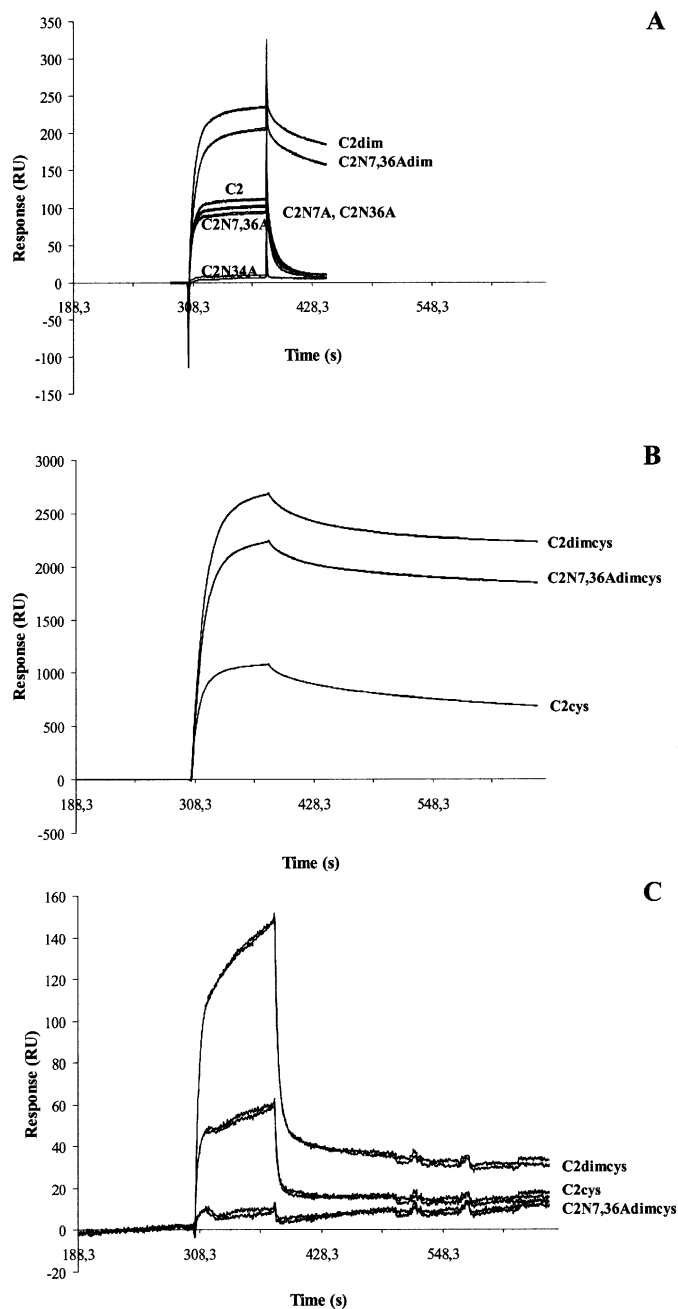
Fig. 2. Biospecific interaction analysis performed on the Biacore™ instrument (Biacore). (**A**) Human polyclonal Fc fragment of IgG is immobilized on the surface and C2 and the different asparagine variants are injected over the surface at a protein concentration of 400 nM. Double samples of each variant are visualized. (**B**) C2 variants are immobilized on the surface using the free thiol of a C-terminal cysteine. Human polyclonal Fc is injected over the surface at a protein concentration of 300 nM. Double samples for each variant are visualized. (**C**) C2 variants are immobilized on the surface using the free thiol of a C-terminal cysteine. Human polyclonal Fab is injected over the surface at a protein concentration of 300 nM. Double samples for each variant are visualized.

Affinity chromatography with an integrated CIP step

The ability of C2 and all mutated variants thereof to function as affinity ligands in an ordinary purification scheme for IgG antibodies was investigated by immobilization of the different domains on HiTrap™ columns (Amersham Biosciences). Between the different cycles, a 2-min CIP step consisting of 0.1 M NaOH was integrated. The column with immobilized
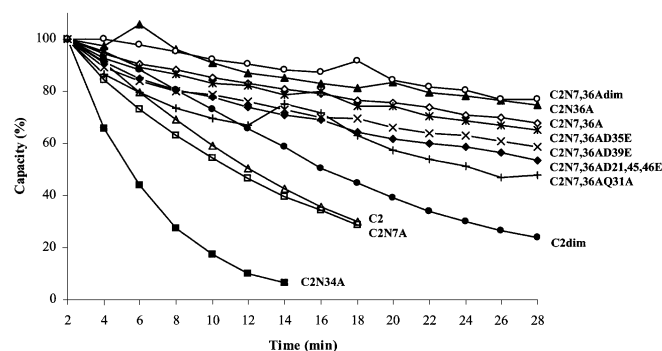
Fig. 3. A comparison of the capacity after repeated CIP treatment following a standard affinity chromatography protocol with C2 and all the mutated variants as immobilized ligands. The cleaning agent is 0.1 M sodium hydroxide.

Table I. Parameters of the interaction of C2 and mutants thereof to Fc of human IgG measured by SPR

	k_a (M^{-1} s^{-1})	k_d (s^{-1})	K_A (M^{-1})
C2	2.9×10^5	27.0×10^{-3}	1.1×10^7
C2$_{N7A}$	2.3×10^5	21.1×10^{-3}	1.1×10^7
C2$_{N36A}$	2.0×10^5	13.2×10^{-3}	1.5×10^7
C2$_{N7,36A}$	2.2×10^5	13.4×10^{-3}	1.7×10^7

All calculations are made with Fc immobilized on the surface. k_a is the association rate constant, k_d is the dissociation rate constant and K_A is the affinity constant (k_a / k_d).

wild-type protein showed a fast decrease in capacity. After a total of 18 min of exposure towards NaOH the C2 column has only ~30% capacity left. The C2$_{N7A}$ column reveals a similar degradation pattern. The C2$_{N34A}$ column loses capacity even faster than wild-type (Figure 3). This column also shows much lower initial binding capacity probably because the Asn34 is involved in the Fc interaction. However, the C2$_{N36A}$ column shows a remarkable increase of stability. After 28 min of exposure this variant still exhibits almost 90% capacity. The double mutant C2$_{N7,36A}$ shows a similar behavior (Figure 3).

In order to find out if further stabilization could be achieved by aspartate replacements, a second generation of mutants was made. In order to be able to detect the potentially small increase in stability these mutations can add, the double mutant, C2$_{N7,36A}$, was used as scaffold. Two single mutants and one triple mutant were made with C2$_{N7,36A}$ as parental molecule, in which the aspartate residues were substituted for glutamate residues, C2$_{N7,36AD35E}$, C2$_{N7,36AD39E}$ and C2$_{N7,36AD21,45,46E}$. Glutamates were chosen in order to retain the charge of the domain. Furthermore, in contrast to aspartate residues no studies have been reported on isomerization of glutamate residues, probably due to the extra –CH$_2$ group of glutamates. Also, the double mutant C2$_{N7,36A}$ was used as scaffold for the single glutamine residue substitution. This single glutamine was substituted for alanine resulting in C2$_{N7,36AQ31A}$. Interestingly, the different aspartate and the glutamine substitutions reveal no further stabilization towards the CIP treatment (Figure 3).

Also, the dimer of the double mutant C2$_{N7,36A}$ exhibited improved stability towards alkaline conditions compared to the wild-type protein, indicating that the linker region used is not degraded during the CIP step. This agrees with the fact

that the linker does not include any asparagines, aspartates or glutamines.

Circular dichroism spectroscopy

The secondary structure content of C2 and the different mutated variants was compared using CD spectroscopy (Johnson, 1990; Schmid, 1997). This technology is suited to detect secondary structure changes in proteins. A protein constituted of all α-helix structure exhibits minima at 208 and 222 nm and a maximum at ~192 nm. Similarly, a protein constituted of β-sheet structure exhibits a minimum at 217 nm and a maximum at ~198 nm (Schmid, 1997). The secondary structure of C2 represents a mixture of α-helix and β-sheet components, but the β-sheet structure predominates. As can be seen from the far-UV CD spectra, all variants show a typical spectrum of a protein containing only a moderate amount of α-helix structure. Hence, all variants exhibit the characteristic parental protein G fold. The CD signal is only slightly decreased for some of the variants compared to wild-type C2, indicating only small changes in the secondary structure upon mutation (Figure 4).

C2$_{N7,36A}$ and the wild-type C2 were also subjected to thermal denaturation to detect change in thermal stability for the mutated variant. Signals were recorded at 217 nm (a typical minimum for a β-sheet protein). The denatured post-transitional state was defined as a random coil structure with a minimum response at 217 nm. Also, the criterion of the pre-transitional state was defined as a folded and fully active domain with a maximum response at 217 nm. Since both proteins were shown to adopt clearly defined pre- and post-transitions, the signals could be transformed to F_{app}, the apparent fraction of unfolded protein. $T_m^{1/2}$ (the midpoint of the denaturation transition) was read from the graph. The C2 domain was found to have a $T_m^{1/2}$ value of 77°C (Figure 5A). The C2$_{N7,36A}$ mutant revealed a slightly decreased thermal stability and its $T_m^{1/2}$ value was set to 71°C (Figure 5A). Reversibility of the denaturation was investigated by allowing the proteins to refold by decreasing the temperature. Both C2 and the mutated variant were able to fully refold again. Accordingly, these proteins seem to adopt a reversible two-state denaturation transition as reported in another study (Alexander *et al.*, 1992a,b).

The chemical stability was also investigated for the wild-type and the C2$_{N7,36A}$ mutant. The criteria for the different structural states were defined as for the thermal denaturation. Signals were recorded at 217 nm and transformed to F_{app} since clearly defined pre- and post-transitions were recorded. As can be revealed from the graph there is no change in structure up to ~2 M GdnHCl but thereafter, there is a progressive change in the CD signal indicating a loss of secondary structure elements (Figure 5B). The $C_m^{1/2}$ value for the C2 domain was read from the graph at F_{app} = 0.5 to be 2.9 M (Figure 5B). This is in agreement with another stability study of the homologous C1 domain (O'Neil *et al.*, 1995). The chemical stability was slightly decreased for the mutated protein. Its $C_m^{1/2}$ value was set to 2.5 M (Figure 5B).

RP-HPLC analysis after NaOH incubation

RP-HPLC analysis was also used to detect degradation of C2 and C2$_{N7,36A}$ during NaOH incubation. After 2 h of incubation in 0.1 M NaOH the original peak of the wild-type protein totally disappeared (data not shown). In contrast to wild-type, the double mutant still shows a peak (slightly decreased) and
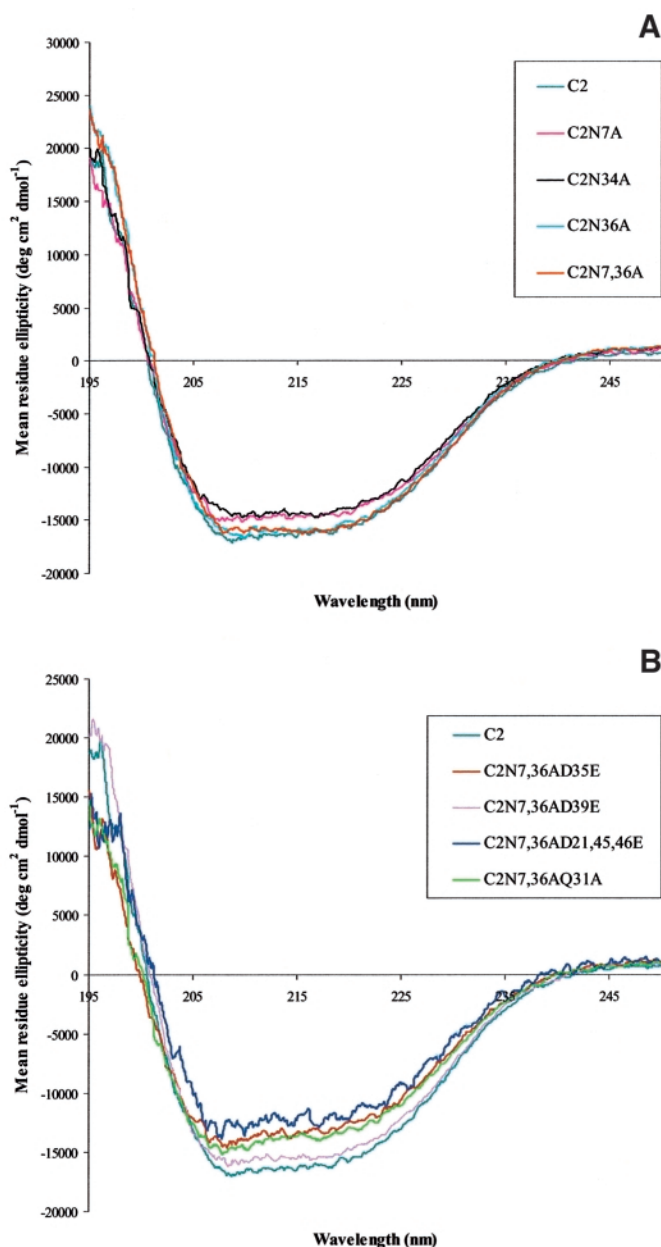
Fig. 4. A CD spectrum visualizing the secondary structure content of all variants constructed. The concentration is 0.1 mg/ml in a phosphate buffer at pH 6 and 20°C. (**A**) C2 and the different variants where asparagine has been replaced. (**B**) C2 and the variants where aspartic acid or glutamine has been exchanged.

Fig. 5. (**A**) Thermal stability analysis of C2 (■) and the mutated variant C2$_{N7,36A}$ (●). The protein concentration is 0.05 mg/ml in an acetate buffer at pH 5.3. Signals at 217 nm have been recorded and transformed to apparent fraction of unfolded protein, F_{app}. (**B**) Chemical stability analysis of C2 (■) and C2$_{N7,36A}$ (●). The protein concentration is 0.25 mg/ml in an acetate buffer at pH 5.3 and different concentrations of GdnHCl at 20°C. Signals at 217 nm have been transformed to F_{app}.

Fig. 6. IEF under non-denaturing conditions. The change in charge for C2 and C2$_{N7,36A}$ is visualized after incubation in 0.1 M sodium hydroxide for 2 h. Lanes 2 and 3, C2 before and after incubation, respectively; lanes 5 and 6, C2$_{N7,36A}$ before and after incubation, respectively; lanes 1, 4 and 7, markers indicating low pH 2.5–6.5, broad pH 3–10 and high pH 5–10.5, respectively.

mass spectrometry analysis reveals the same molecular weight as untreated protein, 7577 Da (data not shown).

Isoelectric focusing gel electrophoresis

The deamidation event is readily detected using IEF since the asparagines and glutamines are modified to aspartate and glutamate residues, respectively. This results in an increased negative charge of the protein. The gel can be run under non-denaturing conditions since all relevant residues are surface exposed. The theoretical p*I* of C2 and C2$_{N7,36A}$ is 4.66 and for each deamidated asparagine or glutamine a decrease of 0.12 p*I* units should be expected. Observed p*I* is 4.55 for untreated C2 and the double mutant (Figure 6). After 2 h of incubation in 0.1 M NaOH the C2 protein reveals a total transformation to a ne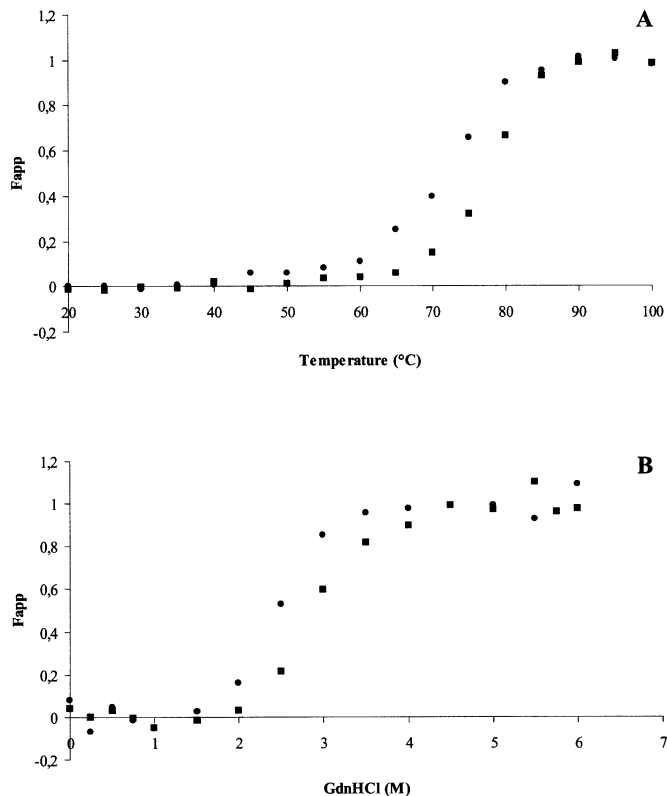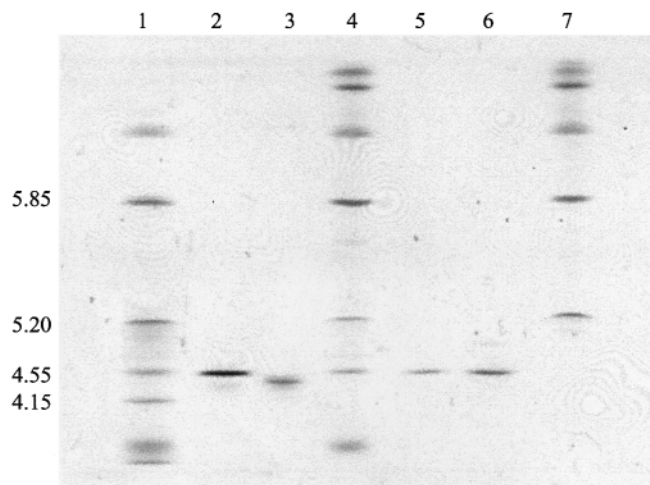w more negatively charged species (Figure 6). The double mutant on the other hand reveals no additional charged species indicating that the deamidation-susceptible residue is substituted (Figure 6).

Discussion

In this study, we have further shown that a simple and straightforward strategy consisting of replacing the asparagine residues can dramatically improve the stability towards alkaline conditions. This protein engineering strategy reduces the sensitivity of protein-based affinity ligands towards alkaline pH often employed in industrial CIP-protocols. In a previous study, ABD, derived from protein G (Kraulis *et al.*, 1996) was stabilized using the same strategy. This domain showed a remarkably increased stability towards NaOH (Gülich *et al.*, 2000b). In the ABD study all asparagines were replaced by other amino acids occurring in homologous sequences, but we emphasized that asparagine replacements have to be considered on a case-by-case basis. Hence, in this study we addressed the fragility of the C2 domain, an IgG-binding domain derived from protein G, by designing single mutants in order to reveal the contribution of each asparagine to the deactivation rate upon exposure towards NaOH.

The biospecific interaction analysis revealed that the two single mutations N7A and N36A resulted in constructs with retained affinity to Fc (Figure 2A; Table I). However, a significant decrease in affinity to IgG was observed when substituting Asn34 (Figure 2A). This is rather predictable since Asn34 is one of the most important amino acids in the binding surface towards Fc. Sloan and Hellinga have also described this (Sloan and Hellinga, 1999). In that particular study, they report a 50-fold increase in K_d for the N34A mutation.

The potential of the single asparagine mutants to function as immobilized ligands during standard industrial procedures was investigated. As revealed from that analysis, Asn36 seems to be the residue most involved in the alkaline sensitivity, since the C2$_{N36A}$ column exhibits a significant increase in stability towards NaOH (Figure 3). An attempt to replace all asparagines in one single mutant was made. Interestingly, this variant was not possible to express in *E.coli* neither as a single domain nor as a genetic fusion to ABD, possibly due to structural instability. Since C2$_{N36A}$ was shown to increase the alkaline stability remarkably, the substitution of that amino acid was most important. However, since Asn7 easily could be exchanged without negative effects on affinity or stability we chose the double mutant, C2$_{N7,36A}$, as scaffold for further stabilization. This double mutant showed retained affinity to polyclonal Fc of human IgG (Figure 2A; Table I) and also improved behavior in the column analysis (Figure 3). In addition to the asparagine replacements, we also replaced the single glutamine residue. Glutamines can also be deamidated by the same mechanism as asparagine residues. However, due to the extra –CH$_2$ group, glutamines are not as frequently deamidated as asparagines. This is a result of the strained energy of a six-membered cyclic imide intermediate compared to the five-membered succinimide intermediate of asparagines. Also, the aspartate residues were replaced using C2$_{N7,36A}$ as scaffold. Aspartate can be modified to its isomer but this modification is not observed as frequently as deamidation of asparagine residues due to loss of a good leaving group at high pH (Stephenson and Clarke, 1989; Tomizawa *et al.*, 1995). Interestingly, according to the NaOH-exposure experiments, none of the different aspartate or glutamine replacements could further increase the alkaline stability (Figure 3). Apparently, none of them are modified in this domain.

The structural and stability analyses indicate that the structure is retained for the mutated proteins with only a slight decrease in secondary structure for some variants (Figure 4). Thus, the loss in affinity for C2$_{N34A}$ occurs without major changes in structure. Also, the double mutant shows slightly decreased thermal stability compared to wild-type but retains the ability to refold after denaturation (Figure 5A). Interestingly, C2$_{N7,36A}$ shows a decreased chemical stability as investigated by the CD analysis (Figure 5B). Alanine has earlier been suggested to be an amino acid that stabilizes α-helical structures (Horovitz *et al.*, 1992; Myers *et al.*, 1997). When substituting Asn34, which is situated in the α-helix, no increased alkaline stability can be detected. Moreover, when substituting Asn36, which is situated in a loop region, for an alanine, an increased stability towards alkaline conditions is detected. It is noteworthy that the structural stability of the selected mutant, C2$_{N7,36A}$, is lower than for the wild-type protein. Despite this fact, this double mutant reveals a major stabilization towards CIP treatment compared to wild-type C2 (Figure 3). This indicates that we have specifically stabilized the domain towards alkaline conditions.

IEF reveals that the double mutant gives no deamidated species in contrast to the wild-type C2, which gives a new, more negatively charged variant (Figure 6). Asn34 is involved in the interaction to Fc, and when substituted, the affinity to IgG is remarkably decreased (Figure 2A). Therefore, this amino acid is not mutated. The single glutamine does not seem to be deamidated when exposed to NaOH. Aspartate residues have been reported to be able to go through an isomerization reaction (Stephenson and Clarke, 1989; Tomizawa *et al.*, 1995). However, this modification cannot be detected on an isoelectric focusing gel. However, the HiTrap™ analyses reveal no further stabilization when substituting the aspartate residues (Figure 3). Accordingly, these residues are not involved in the sensitivity towards alkaline environment of this domain. However, there might be amino acids in the structure that go through an isomerization or deamidation event that do not remarkably affect the activity of the C2 domain.

Finally, genetically fused dimers of C2 and C2$_{N7,36A}$ result in increased binding capacity according to the directed immobilization using the C-terminal cysteine (Figure 2B). Also, the linker region connecting the individual monomers is stable towards NaOH treatment (Figure 3).

Asn36 is situated in the loop connecting the α-helix with the third β-strand. Numerous studies have reported on the dependence on the sequence and conformation for deamidation (Kossiakoff, 1988; Lura and Schirch, 1988; Wearne and Creighton, 1989). A significant effort has been dedicated to screen the susceptibility of different sequences (Robinson and Robinson, 1991, 2001). The most crucial factor is the C-terminal residue succeeding asparagine. The most sensitive sequence is known to be Asn–Gly. The high deamidation rate for this sequence is due to the small side chain of glycine, which introduces a high flexibility in the backbone of the domain. When a glycine succeeds an asparagine, the main peptide chain and the side chain torsional angles are favorable for the deamidation mechanism. Moreover, these angles occur almost exclusively in the more flexible loop regions (Wright, 1991; Xie and Schowen, 1999; Xie *et al.*, 2000). Asn36 precedes a glycine residue and is also situated in a loop (Figure 1). Both these parameters are probably affecting the susceptibility towards deamidation. Asn7 is also followed by a glycine residue but does not seem to be deamidated (Figure 1). A probable explanation for this is that the rigidity of the β-sheet structure may restrain the flexibility of the individual

residues. Asn34 is situated in the α-helix and an aspartate is flanking C-terminal (Figure 1). Residues in α-helices are less frequently deamidated since the torsional angles are not optimal for the reaction mechanism (Kosky *et al.*, 1999).

In conclusion, we have stabilized one of the IgG-binding domains of streptococcal protein G towards alkaline treatment. Our protein engineering strategy involves replacement of asparagine residues in order to increase the alkaline stability of a protein. However, each protein needs to be considered on a case-by-case basis since all asparagines may not be possible to replace due to activity losses or structural destabilization. Obviously, some asparagines are not participating in the deactivation process due to structural and conformational parameters. By designing single mutants the susceptible residue can be revealed. With this study we have been able to successfully stabilize the IgG-binding region C2 of protein G and can present a prototype for an improved protein G affinity chromatography medium. Such improvements will be crucial to obtain industrial implementation of protein-based affinity ligands, which normally show too high a susceptibility towards the alkaline conditions used during regeneration of the chromatography devices. This protein engineering strategy might make SPG-based affinity chromatography media useful for cost-efficient industrial-scale purification of monoclonal antibodies for therapeutic use.

Acknowledgements

References

Achari,A., Hale,S.P., Howard,A.J., Clore,G.M., Gronenborn,A.M., Hardman,K.D. and Whitlow,M. (1992) *Biochemistry*, **31**, 10449–10457.

Åkerström,B., Brodin,T., Reis,K. and Björck,L. (1985) *J. Immunol.*, **135**, 2589–2592.

Alexander,P., Fahnestock,S., Lee,T., Orban,J. and Bryan,P. (1992a) *Biochemistry*, **31**, 3597–3603.

Alexander,P., Orban,J. and Bryan,P. (1992b) *Biochemistry*, **31**, 7243–7248.

Asplund,M., Ramberg,M. and Johansson,B.-L. (2000) *Proc. Biochem.*, **35**, 1111–1118.

Björck,L. and Kronvall,G. (1984) *J. Immunol.*, **133**, 969–974.

Derrick,J.P. and Wigley,D.B. (1992) *Nature*, **359**, 752–754.

Derrick,J.P. and Wigley,D.B. (1994) *J. Mol. Biol.*, **243**, 906–918.

Fassina,G., Verdoliva,A., Odierna,M.R., Ruvo,M. and Cassini,G. (1996) *J. Mol. Recognit.*, **9**, 564–569.

Ford,C.F., Suominen,I. and Glatz,C.E. (1991) *Protein Expr. Purif.*, **2**, 95–107.

Gallagher,T., Alexander,P., Bryan,P. and Gilliland,G.L. (1994) *Biochemistry*, **33**, 4721–4729.

Geiger,T. and Clarke,S. (1987) *J. Biol. Chem.*, **262**, 785–794.

Gill,S.C. and von Hippel,P.H. (1989) *Anal. Biochem.,* **182**, 319–326.

Goward,C.R., Irons,L.I., Murphy,J.P. and Atkinson,T. (1991) *Biochem. J.*, **274**, 503–507.

Gräslund,T., Lundin,G., Uhlén,M., Nygren,P.-Å. and Hober,S. (2000) *Protein Eng.*, **13**, 703–709.

Gronenborn,A.M., Filpula,D.R., Essig,N.Z., Achari,A., Whitlow,M., Wingfield,P.T. and Clore,G.M. (1991) *Science*, **253**, 657–661.

Gülich,S., Uhlén,M. and Hober,S. (2000a) *J. Biotechnol.*, **76**, 233–244.

Gülich,S., Linhult,M., Nygren,P.-Å., Uhlén,M. and Hober,S. (2000b) *J. Biotechnol.*, **80**, 169–178.

Higuchi,R., Krummel,B. and Saiki,R.K. (1988) *Nucleic Acids Res.*, **16**, 7351–7367.

Horovitz,A., Matthews,J.M. and Fersht,A.R. (1992) *J. Mol. Biol.*, **227**, 560–568.

Jansson,M., Li,Y-C., Jendeberg,L., Anderson,S., Montelione,G.T. and Nilsson,B. (1996) *J. Biomol. NMR*, **7**, 131–141.

Johnson,C.W. (1990) *Proteins: Struct. Funct. Genet.*, **7**, 205–214.

Kosky,A.A., Razzaq,U.O., Treuheit,M.J. and Brems,D.N. (1999) *Protein Sci.*, **8**, 2519–2523.

Kossiakoff,A.A. (1988) *Science*, **240**, 191–194.

Kraulis,P.J., Jonasson,P., Nygren,P.-Å., Uhlén,M., Jendeberg,L., Nilsson,B. and Kördel,J. (1996) *FEBS Lett.*, **378**, 190–194.

Kuszewski,J., Gronenborn,A.M. and Clore,G.M. (1999) *J. Am. Chem. Soc.*, **1**, 2337–1999.

Li,R., Dowd,V., Stewart,D.J., Burton,S.J. and Lowe,C.R. (1998) *Nat. Biotechnol.*, **16**, 190–195.

Lian,L.-Y., Derrick,J.P., Sutcliffe,M.J., Yang,J.C. and Roberts,G.C.K. (1992) *J. Mol. Biol.*, **228**, 1219–1234.

Lura,R. and Schirch,V. (1988) *Biochemistry*, **27**, 7671–7677.

Myers,J.K., Pace,N. and Scholtz,J.M. (1997) *Biochemistry*, **36**, 10923–10929.

Nygren,P.-Å., Eliasson,M., Abrahmsén,L., Uhlén,M. and Palmcrantz,E. (1988) *J. Mol. Recognit.*, **1**, 69–74.

Olsson,M.O. and Isaksson,L.A. (1979) *Mol. Gen. Genet.*, **169**, 251–257.

Olsson,A., Eliasson,M., Guss,B., Nilsson,B., Hellman,U., Lindberg,M. and Uhlén,M. (1987) *Eur. J. Biochem.*, **168**, 319–324.

O'Neil,K.T., Hoess,R.H., Raleigh,D.P. and DeGrado,W.F. (1995) *Proteins: Struct. Funct. Genet.*, **21**, 11–21.

Robinson,A.B. and Robinson,L.R. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 8880–8884.

Robinson,N.E. and Robinson,A.B. (2001) *Proc. Natl Acad. Sci. USA*, **98**, 944–949.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.

Sauer-Eriksson,A.E., Kleywegt,G.J., Uhlén,M. and Jones,T.A. (1995) *Structure*, **3**, 265–278.

Sayle,R.A. and Milner-White,E.J. (1995) *Trends Biochem. Sci.*, **20**, 374.

Schmid,F.X. (1997) In Creighton,T.E. (ed.), *Protein Structure*. Oxford University Press, Oxford, pp. 261–297.

Sloan,D.J. and Hellinga,H.W. (1999) *Protein Sci.*, **8**, 1643–1648.

Ståhl,S., Nilsson,J., Hober,S., Uhlén,M. and Nygren,P.-Å. (1999) In Fleckinger,M.C. and Drew,S.W. (eds), *The Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis and Bioseparation*. John Wiley and Sons Inc., New York, pp. 49–63.

Stephenson,R.C. and Clarke,S. (1989) *J. Biol. Chem.*, **264**, 6164–6170.

Tomizawa,H., Yamada,H., Hashimoto,Y. and Imoto,T. (1995) *Protein Eng.*, **8**, 1023–1028.

Wearne,S.J. and Creighton,T.E. (1989) *Proteins: Struct. Funct. Genet.*, **5**, 8–12.

Wright,H.T. (1991) *Crit. Rev. Biochem. Mol. Biol.*, **26**, 1–52.

Xie,M. and Schowen,R.L. (1999) *J. Pharm. Sci.*, **88**, 8–13.

Xie,M., Aubé,J., Borchardt,R.T., Morton,M., Topp,E.M., Vander Velde,D. and Schowen,R.L. (2000) *J. Peptide Res.*, **56**, 165–171.