Kinetic and affinity predictions of a protein-protein interaction using multivariate experimental design *

Erwin De Genst‡¶, Daphne Areskoug§, Klaas Decanniere‡, Serge Muyldermans‡ and Karl Andersson§

* This work was supported by the Vlaams Interuniversitair Instituut voor Biotechnologie and the Fonds voor Wetenschappelijk Onderzoek Vlaanderen

‡ Dienst Ultrastructuur, Vlaams Interuniversitair Instituut Biotechnologie, Vrije Universiteit Brussel, Paardenstraat 65, B-1640 Sint-Genesius Rode, Belgium

§ Biacore AB, Rapsgatan 7, SE-754 50 Uppsala, Sweden

¶ To whom all correspondence should be addressed. Dienst Ultrastructuur, Vlaams Interuniversitair Instituut Biotechnologie, Vrije Universiteit Brussel, Paardenstraat 65, B-1640 Sint-Genesius Rode, Belgium E-mail: edegenst@vub.ac.be, tel. 32-2-359-02-82; Fax: 32-2-359-02-89

Running Title: Kinetic and affinity predictions of protein-protein interactions
Abstract

We measured the influence of 14 mutations and 5 environmental variables (buffer perturbation) on the association and dissociation rate of a camel single domain antibody (cAb-Lys3) interacting with hen egg-white lysozyme using a surface plasmon resonance based biosensor. Based on this data set we constructed quantitative predictive models for both kinetic ($k_a$ and $k_d$) constants as for the affinity constant ($K_d$). Mutations, after parameterization by quantitative descriptors, and buffers were selected using multivariate experimental design. These models were able to predict the corresponding parameters of four new variants of cAb-Lys3. Moreover the models provide insights to the important chemical aspects of the interacting residues, which are difficult to deduce from the crystal structure. Our approach provides useful physico-chemical information of protein-protein interactions in general. The information obtained from this kind of analysis complements and goes beyond that of conventional methods like alanine scanning and substitution by closely related amino acids. The mathematical modeling may contribute to a rational approach in the optimization of bio-molecules of biotechnological interest.

Introduction

Protein-protein interactions attract a lot of attention in this era of proteomics and antibody technology. Naive and synthetic antibody libraries in combination with phage and ribosome display are now routinely used to obtain antibody molecules with desired properties and specificities (1-6). However, the affinity and/or stability of the original binder retrieved from the library often requires improvements. Molecular evolution techniques (1-8) and rational design using sophisticated computer programs (9-16) are two approaches to achieve this goal. Drawbacks to the former strategy are its dependence on the generation of large libraries, the
difficulty in discrimination between binders with small differences in affinity, and the bias in favor of binders with improved expression levels (1;3). Also codon bias and the generation of non-functional or deleterious mutants are major drawbacks of this strategy (1;3). The rational design strategy largely depends on the availability of the structure of the protein-protein complex and although predictive algorithms have been improved significantly over the years, predictions of interaction energies, especially those related to amino acid modifications, remain highly problematic (9-16).

It is generally accepted that, experimental data on the energetic contribution of the modified amino acids is required to develop algorithms relating binding energetics to particular amino acid modifications. Many analyses have used site directed mutagenesis, especially alanine-scanning (17-21) and substitution by closely related amino acids (21-23), to determine the energetic contributions of the amino acid side chains and functional groups. These strategies have shown to be very attractive thanks to their simplicity. However, they often lead to misinterpretations and inconclusive results (21-25). One of the possible reasons for these failures is that quantitative measures like kinetic and affinity constants have been related to amino acid substitutions in a qualitative or intuitive way, effectively reducing the analysis to the presence or absence of a side chain or functional group. Other chemical property differences between amino acids are not considered. For instance, a tyrosine residue is often considered to have the same chemical nature as a phenylalanine, except for an additional hydroxyl group (19;21;23). It is not taken into account that addition or removal of functional groups of a residue may influence the chemical properties of the residue as a whole and thereby changing the chemical nature of the interaction more than intended. To analyze the effect of various physico-chemical properties of the amino acids, replacement of a residue by a range of different amino acids may be required. Also, the
introduced modification has to be quantified in order to extract a mathematical correlation model. ‘Quantitative structure activity relationship’ (QSAR⁴) is a method of choice to construct such models and have already been used extensively in the field of rational drug design. The QSAR uses descriptive scales (e.g. electronic, steric and hydrophobic properties) to account for the observed correlation between the structure and activity of protein-ligand interactions (26-32). The QSAR models are predictive and are often combined with powerful multivariate statistical methods, hereby extracting a maximum of information from a minimum amount of data. In addition, the reliability of the data and the models are estimated (26;30).

In this paper, we demonstrate the use of QSAR to investigate the interaction of the camel single domain antibody, cAb-Lys3, with its antigen, hen egg-white lysozyme (33;34). The cAb-Lys3 antibody fragment is a well studied model system, of which several crystal structures in complex with lysozyme are available (33-35).

Two residues of the CDR3 loop of cAb-Lys3 were substituted simultaneously to various amino acids and quantitative descriptions for the substituted residues were used. In our analysis we related the sequence of these two positions to the kinetic parameters of the interaction. In addition, the binding of all mutants to lysozyme was characterized in several buffers containing a variety of chemical additives, to explore the sensitivity of binding due to changes in chemical environment for each mutant. Both, the amino acid replacements and the buffer additives were selected using a multivariate experimental design.

Materials & Methods

Reagents. All reagents were analytical grade. Hen egg-white lysozyme (HEL) was purchased from Roche chemicals. PDEA, EDC, NHS, ethanolamine and DTE were obtained from Biacore
AB and cystamine was obtained from Sigma Aldrich. Restriction enzymes were purchased from Life Technologies Gibco BRL.

Amino acid numbering and notation. The amino acids are numbered as they occur in the linear sequence of cAb-Lys3. Positions 101 and 105 correspond to respectively 97 and 100a in the Kabat numbering (36). The cAb-Lys3 mutants are referred to by a two-letter code where the first and second letter represents the amino acid (single letter code) at position 101 and 105, respectively. The wild type cAb-Lys3 is thus represented by TS because it possesses threonine and serine at the respective positions.

Selection of the positions for modification. We need to identify positions of cAb-Lys3 where an amino acid substitution would modulate antigen binding but does not abolish the interaction. These are to be found at the periphery of the interaction interface (37). We analyzed the crystal structure data of the cAb-Lys3 with HEL (34) and looked for (i) two amino acids (ii) within the CDR3-loop, (iii) that were not in direct contact with each other, (iv) and that are at the edge of the antibody/antigen interaction surface. The residues at positions 101 and 105 comply with those criteria. In addition, we were attracted by the Ser105, because it is located at the tip of the CDR3 loop of cAb-Lys3 and it is in contact with the catalytic residues Glu35 and Asp52 of HEL (34). Furthermore, we knew from previous (unrelated) experiments that Ser105 could be substituted without dramatic effects on the cAb-Lys3 affinity for HEL (M. Murshed, E. De Genst, unpublished results).
Experimental design of the amino acid replacements. In the selection of the amino acid substitutions at positions 101 and 105 of cAb-Lys3, we excluded tryptophane, tyrosine and phenylalanine at either position since inspection of the crystal structure revealed that these residues at those positions would lead to large sterical clashes with the antigen. Also cysteine was not included to avoid dimerization and folding problems of the recombinant mutant proteins. The amino acids at positions 101 and 105 were parameterized using three scales, ZZ1, ZZ2 and ZZ3, representing respectively hydrophilicity, size/polarizability and electronic properties (polarity, electrophilicity, charge, electronegativity) (27) (table 1). In order to resolve 6 parameters in the model, more than 6 mutants are required. We designed 8 new mutants, mainly double mutants, and added 6 already available single mutants (TG, TA, TP, TH, TQ, TN). Therefore, 14 single and double mutants at position 101 and 105 of cAb-Lys3 were envisaged. The final selection of the 8 new mutants was based on the optimization of the condition number of the design matrix (29;38). The condition number is a measure of the degree of linear dependency between the rows of the descriptor-matrix. For a statistically perfect design, the condition number would be 1 (29). Condition numbers between 1 and 5 are considered to be good, above 5 the design becomes progressively less acceptable (29). The optimization of the condition number was performed by manually testing various combinations of double amino acid mutations (positions 101 & 105) combined with the preset single mutants (position 105), until a condition number of 2.81 was obtained. An ordinary D-optimal design generator could not be applied here, since this algorithm assumes that all descriptors can be varied independently, which is not the case for amino acids (three parameters are fixed per amino acid).
Generation of the cAb-Lys3 mutants. Single mutants TG, TA, TP, TM, TH, TQ, TN were initially constructed by randomization of the codon at position 105 in a cAb-Lys3 pHEN4 vector (39) by PCR. The sequence of the single mutants was analyzed on an ALF automatic sequencer (Amersham Biosciences) and recloned in a vector in which the hemaglutinin tag and the geneIII between NotI and EcoRI sites were replaced by a fragment encoding six consecutive histidine codons followed by a termination codon. This modified expression vector is called pHEN6 cAb-Lys3 (39). The mutants LS, VN, RT, PG, SQ, HD, QP and MV of cAb-Lys3 were constructed in pHEN6 cAb-Lys3 vector by site directed mutagenesis of closed circular DNA in vitro, according to Chen et al. (40). Primers were 53 nucleotides in length and are complementary to the coding strand of the cAb-Lys3 gene. The sequence of the primers were:

5’-CAG ACC GTG GCC GCA TTC ATA ATA XXX AGC GTA GAT X

CGA ATC TGC CGC AC-3’, where XXX is the mutagenic codon. All of these mutants except MV removed a unique TaqI (underlined) site in the original sequence of cAb-Lys3 and mutants were screened by colony PCR amplifying the cAb-Lys3 gene and digestion with TaqI. The MV mutant introduced a Sau3AI (dotted underlined) site and MV clones were screened by digestion with Sau3AI. Mutants were confirmed by DNA sequencing. The single mutants SS and PS were intended to encode SQ and PG respectively.

Expression & purification. Periplasmic expression, extraction and purification by immobilized metal affinity chromatography (IMAC) and gelfiltration of mutants and wild type cAb-Lys3 were performed according to Lauwereys et al. (39). The concentration of the mutants and wild type cAb-Lys3 was determined spectrophotometrically at 280nm using the calculated sequence based extinction coefficient (41) which is equal to 27220 M⁻¹cm⁻¹ for all mutants.
Immobilization of PDEA modified lysozyme. Lysozyme was modified with PDEA to introduce reactive disulphide groups before immobilization via surface thiol coupling, as described in BIAapplications Handbook (42). This immobilization method was chosen based on the closest approximation to a 1:1 binding model for the binding of 8 mutants and the wild-type cAb-Lys3 to the immobilized ligand.

PDEA modified lysozyme dissolved in 10mM NaAc buffer pH5.5 was coupled to CM5 chips at a rate of 5 µl/min. Two flow cells (fc1 and fc3) were incubated with 35 µl EDC/NHS and subsequently with 35 µl 1M ethanolamine pH 8.5, and served as reference surfaces for the kinetic measurements. Two other flow cells (fc2 and fc4) were used to immobilize PDEA-modified lysozyme at 80 RU and 500 RU respectively. The two flow cells were treated with 10 µl EDC/NHS and subsequently with 15 µl cystamine dihydrochloride, followed by 15µl DTE. Manual injection of the PDEA modified lysozyme was then performed until the aimed immobilization levels were obtained. Excess reactive thiols on the chip were deactivated by 20µl PDEA-NaCl.

Experimental design of the perturbation buffers. All buffers used were phosphate buffered saline (PBS) based. The chemical additives NaCl, urea, DMSO, KSCN and pH, and their ranges were chosen as in Andersson et al. and Choulier et al. (29;38). To analyze the main effects of these five chemical factors on the (mutant) antibody antigen interaction, a standard fractional factorial design to a resolution of III was applied, which consists of 11 experiments (2^{5-2} buffers to probe the high and low value of five factors without resolving interaction effects, plus 3 replicate centerpoints ) (see table 2) (43).
Measurements/fittings/calculations. Four Biacore® 3000 instruments and four Sensor Chip CM5 were used to obtain the data at 25°C and 30µl/min, with PBS supplemented with 3mM EDTA and 0.005% surfactant P20 (table 2) as running buffer. Replicate runs were performed on a different chip and instrument, which made the measurements independent. Each mutant was characterized in PBS at four protein concentrations ranging from 12µM to 31.25nM, depending on the affinity of the mutant for HEL and was subsequently tested in 11 perturbation buffers (in randomized run order). The mutants were injected for 2 minutes and dissociation data was collected during 5 minutes (perturbation buffers), or 10 minutes (PBS). Regeneration after each cycle was performed using 10mM glycine-HCl pH 1.0 for 2 minutes.

Control injections using 1500nM of the TG mutant in PBS was performed before and after two kinetic cycles to confirm that the activity of the immobilized lysozyme did not change over time. Zero concentration data (injection of buffer alone) was always subtracted from the sensorgrams before fitting. Kinetic parameters were obtained by global fitting of the sensorgrams to a 1:1 model using BIAevaluation 3.1 software. The Rmax parameter, representing the binding capacity of the chip, was fixed to the average equilibrium response of the TG injections.

Mathematical modeling. All mathematical modeling was performed in Modde 5.0 (Umetrics AB, Umea, Sweden) essentially according to Andersson et al. and Choulier et al. (29;38). A ‘quantitative sequence kinetic relationship’ (QSKR) (29;38) mathematical model relating the logarithm of ka, kd and Kd in PBS buffer to the ZZ scales at position 101 and 105 of the mutants in the design was obtained using multiple linear regression (MLR). Logarithmic transformation
of the kinetic and affinity constants was required to get normally distributed data (a prerequisite for MLR) (27;29).

A ‘quantitative buffer kinetic relationship’ (QBKR) (29;38) model relating the logarithm of $k_a$, $k_d$ and $K_d$ of the mutants in the design to the buffer composition was performed for each replicate separately using MLR. Again logarithmic transformation of the kinetic constants and affinity constant was necessary to obtain normally distributed data. For each QBKR, a sensitivity fingerprint was derived according to Andersson et al. and Choulier et al. (29;38) by normalizing the coefficients (dividing the coefficients by the constant term in the linear model). Quantitative sequence perturbation relationship (QSPR) was performed by relating the sequence descriptors (ZZ scales) of the mutants with the chemical sensitivity fingerprints using partial least squares (PLS).

The quality of all the mathematical models was estimated using leave-one-out cross-validation ($Q^2$) and the correlation coefficient ($R^2$). $Q^2$ reflects the fraction of the variance of the data that can be predicted by the model, as opposed to the correlation coefficient $R^2$, which reflects the percentage of the variance in the responses that can be explained by the model. Large $Q^2$ ($\geq 0.7$), indicates that the model has good predictive ability and will have small prediction errors (44).

Results

Properties of the target positions. The crystal structure of cAb-Lys3 in complex with HEL revealed that Thr101 and Ser105 are at the edge of the interaction interface (34). These residues are part of the protruding CDR3 loop of the cAb-Lys3, and their atoms are not in direct contact with each other in the folded protein (fig. 1). The residues are contacting lysozyme via their side chains. The methyl group of Thr101 fills a hydrophobic cavity formed by residues Asn103 and
Ala107 of lysozyme and the aromatic ring atoms of Tyr103 of cAb-Lys3. The polar hydroxyl group of Thr101 is pointing towards the solvent (fig.2 A). Ser105 is located at the tip of the CDR3 loop of cAb-Lys3 and contacts Gln52, Glu35 and Asp52 of HEL (fig.2 B). With the possible exception of large aromatic side chains, mutants can be generated at these positions that will have a minimal disturbance of the overall antibody-antigen contact. Thus, the cAb-Lys3 amino acids at these positions (101 & 105) were chosen to be substituted. The mutants are expected to affect binding as opposed to abolishment. Indeed, previous affinity measurements of the single mutants at position 105 revealed that Ser105 in the wild-type cAb-Lys3 is not crucial for binding, although the affinity can severely be influenced (see for example table 4 a,b with TN, TH, TQ, TA, TG, TP & TM having a $K_d$ ranging from 700 to 9 nM).

**Experimental design of the amino acid replacements.** The selection of the mutants at positions 101 and 105 was performed as described under Materials and Methods (‘Experimental design of the amino acid replacements’). The result of the experimental design is a matrix of 14 mutants x 6 ZZ scales (table 3 a) with condition number 2.81. The double mutants SS and PS (generated fortuitously by an artefactual chemical synthesis of the mutagenic primer intended to encode SQ and PG respectively), the mutant TM and the wild type cAb-Lys3 (TS) (table 3 b) were not included in the design, since they were not necessary to improve the condition number.

**QSKR or quantitative sequence kinetic relationship.** All kinetic and affinity determinations in PBS buffer are listed in table 4 a, b. The mutants cover a broad range in $k_a$, $k_d$ and $K_d$ (1000 – 100 000 M$^{-1}$s$^{-1}$, 0.0004 – 0.3 s$^{-1}$ and 20 000 – 10 nM, respectively).

11
The values of the kinetic constants and affinity constant were highly reproducible for the replicate runs as well as between the two immobilized flow cells on the same sensor chip. No mass transfer limitations were detected. All binding traces fitted well to the 1:1 binding model.

The QSKR analysis resulted in models with very good statistical properties: \((R^2, Q^2)\) of \((0.93, 0.88), (0.83, 0.85), (0.85, 0.77)\) for \(\log k_a, \log k_d\) and \(\log K_d\) respectively and they are described by the following mathematical equations:

\[
\log k_a = 3.60 - 0.0016 \cdot ZZ^{101} - 0.25 \cdot ZZ^{2101} - 0.18 \cdot ZZ^{3101} + 0.16 \cdot ZZ^{1105} - 0.097 \cdot ZZ^{2105} - 0.19 \cdot ZZ^{3105} \\
\log k_d = -2.0 + 0.051 \cdot ZZ^{101} + 0.092 \cdot ZZ^{2101} + 0.23 \cdot ZZ^{3101} + 0.091 \cdot ZZ^{1105} + 0.0095 \cdot ZZ^{2105} - 0.57 \cdot ZZ^{3105} \\
\log K_d = -5.61 + 0.053 \cdot ZZ^{101} + 0.35 \cdot ZZ^{2101} + 0.41 \cdot ZZ^{3101} - 0.071 \cdot ZZ^{1105} + 0.11 \cdot ZZ^{2105} - 0.38 \cdot ZZ^{3105}
\]

Equations [1], [2] and [3] were used to predict the values of \(k_a, k_d\) and \(K_d\) for three additional mutants of cAb-Lys3 not included in the design as well as for the wild type cAb-Lys3 (table 3 b, table 4 b). Figure 3 A, B, C shows the observed versus predicted values for \(\log k_a, \log k_d\) and \(\log K_d\) for all cAb-Lys3 variants. The plot shows good predictability for all three constants. From figure 3 we also see that the absolute values of \(k_a, k_d\) and \(K_d\) of the four extra proteins are predicted by their models within an order of magnitude.

Inspection of the residuals versus predicted responses (fig. 3 D, E, F) shows that the fit to these linear models is very good, that there are no important interaction terms between the amino acids at the two mutated positions and that there are no important interactions between the ZZ scales per position. This is also reflected in the two double mutant cycles present in the data, namely SQ-TQ-TS-SS-SQ and PG-TG-TS-PS-PG. They showed no significant or small coupling energy.
for all three binding parameters (0.30 \((k_a)\), 0.25 \((k_d)\), -0.04 \((K_d)\) and 0.12 \((k_a)\), -0.53 \((k_d)\), -0.65 \((K_d)\) kcal/mole).

Plots of scaled and centered coefficients of the models are presented in figure 4. We observe that both positions affect the affinity and kinetics of the interaction. The largest negative effects on 
log \(k_a\) are ZZ2\(^{101}\), ZZ3\(^{105}\), ZZ3\(^{101}\), ZZ2\(^{105}\) and ZZ1\(^{101}\) (in descending order). ZZ1\(^{105}\) has a quite significant positive effect on log \(k_a\).

The large negative effect of ZZ3\(^{105}\) is a striking feature in the log \(k_d\) model. ZZ3\(^{101}\), ZZ2\(^{101}\), ZZ1\(^{105}\), ZZ1\(^{101}\) and ZZ2\(^{105}\) in descending order of importance determine positive effects.

The log \(K_d\) model represented by the subtraction of the log \(k_d\) from the log \(k_a\) model, reflects the global binding energy of the interaction as a function of the factors under consideration. In descending order of importance, ZZ3\(^{105}\) and ZZ1\(^{105}\) are the negative factors and ZZ3\(^{101}\), ZZ2\(^{101}\), ZZ2\(^{105}\) and ZZ1\(^{101}\) are the positive factors of the log \(K_d\) model (fig. 4 C).

By comparing the importance of each ZZ scale (fig. 4) to the physico-chemical properties that they represent (table 1) we find that, at position 101, small hydrophobic/apolar amino acids are optimal for all three parameters of binding. At position 105 we observe that a small polar and electrophilic amino acid would be optimal for \(k_d\) and \(K_d\) whereas a fast on-rate \((k_a)\) would be favoured by a small apolar (and non-electrophilic) amino acid.

**QBKR or quantitative buffer kinetic relationship.** The QSKR analysis uses sequence modifications to perturb the protein-protein interaction, to extract information on the physico-chemical characteristics of the interaction. However, the QSKR perturbation is restricted in this case, by the fact that only 19 mutations can be constructed per residue position. Alternatively, perturbation of the interaction can also be achieved by changing environmental parameters.
Therefore, we extended our analysis by measuring the binding kinetics of all our mutants in a number of perturbation buffers. The data for each of the 18 cAb-Lys3 variants (14 designed mutants, TM, SS, PS and wild type cAb-Lys3 (TS)), in this analysis consist of the kinetic and affinity constants measured in 11 perturbation buffers containing five buffer additives, NaCl, urea, KSCN, DMSO and pH.

The values of the kinetic constants and affinity constant were highly reproducible for the replicate runs as well as between the two immobilized flow cells on the same sensor chip. No mass transfer limitations were detected. All binding traces fitted well to the 1:1 binding model. Due to very fast kinetics, no data for $k_a$ and $k_d$ could be obtained for the binding of the SQ mutant in buffers 1, 3, 6, 9 and 11. For these runs $K_d$ data was obtained from the equilibrium responses.

For most mutants we obtained very good statistical properties for the QBKR models. Some mutants (e.g. SQ and TH) resulted in models of poor quality due to high or low $k_d$ values which were barely perturbed by the buffer additives. This data was therefore excluded during further analysis.

Coefficient plots for log $k_a$, log $k_d$ and log $K_d$ for the wild-type cAb-Lys3 are shown in figure 5. The sensitivity fingerprints for log $K_d$ are given for both replicates of each mutant in table 5. (The sensitivity fingerprints for log $k_a$ and log $k_d$ can be obtained from the corresponding author upon request). From table 5 we observe that the normalized coefficients for KSCN of the TQ, TH and RT mutants differ in sign between the replicates because the QBKR coefficient for KSCN falls within the noise level.
**QSPR or quantitative sequence perturbation relationship.** In this analysis we correlated the sensitivity fingerprints of the different mutants to the ZZ scale matrix. The result is a sensitivity model for the kinetic constants for a given chemical substance related to the ZZ scales at position 101 and 105. For example, the sensitivity of log $k_a$ to changes in NaCl concentration is represented by the following model,

$$NaCl_{Sens}k_a = 0.028 - 0.00044 \cdot ZZ1^{101} + 0.0046 \cdot ZZ2^{101} + 0.0022 \cdot ZZ3^{101} - 0.0029 \cdot ZZ1^{105} + 0.0015 \cdot ZZ2^{105} + 0.0029 \cdot ZZ3^{105}$$  \[4\]

In contrast to the coefficients in the QSKR & QBKR models that are directly proportional to the binding parameter, the coefficients in the QSPR model are proportional to the variation of the binding parameter. Depending on the sign (+/-) of the coefficients in the sensitivity fingerprint, the interaction parameter increases or decreases in magnitude upon increasing levels of the added chemical. Correlation of the sensitivity fingerprint and the ZZ scale matrix was obtained by PLS. Only the sensitivity model for NaCl yielded $Q^2$ values above 0.5 and we therefore limit our analysis to these models. The NaCl sensitivity models are called $NaCl_{Sens}k_a$, $NaCl_{Sens}k_d$ and $NaCl_{Sens}K_d$ for $k_a$, $k_d$ and $K_d$ respectively and yielded a $Q^2$ of 0.72, 0.52 and 0.68.

From figure 6 we see that $ZZ2^{101}$ has a large positive effect on the salt sensitivity for $K_d$ whereas $ZZ2^{105}$ has the largest negative effect. This salt sensitivity is predominantly manifested in the $k_a$ parameter for position 101 and in the $k_d$ parameter at position 105.

The interaction of cAb-Lys3 with lysozyme will have a larger decrease in $K_d$ when the salt concentration increases if one or both of the following two conditions are fulfilled: (1) Large hydrophobic and electrophilic amino acids at position 101 and (2) a small hydrophilic amino acid at position 105.
Discussion

Quantitative sequence kinetic relationship (QSKR). Affinity maturation in vitro is a major challenge in antibody engineering. Binders from naive or synthetic scFv libraries with the required antigen specificity, often have sub-optimal affinities. Molecular evolution techniques have been employed as a strategy to improve the affinities of moderate binders (1-8). However, these methods are based on the random insertion of mutations and reselection, so that a favorable mutation might be cancelled by a negative mutation elsewhere in the sequence. Therefore, rational design of modifications within the sequence of these moderate binders constitutes on a longer term a preferred approach to optimize the affinity (9-16). Although many attempts have already been made to predict the effect of amino acid modifications on the affinity, these predictions remain highly problematic and the molecular evolution techniques remain largely the methods of choice. In the present study we aimed at a rational approach to probe two interacting residues in a protein-protein interface and to extract a model which quantifies the contribution of three descriptor variables, ZZ1-3, to the interaction parameters, $k_a$, $k_d$ and $K_d$, measured in a standard buffer. The cAb-Lys3/lysozyme complex was taken and positions 101 and 105 of cAb-Lys3 were chosen to be modified. These residues are located at the edge of the interface and interact with lysozyme via their side chains. The residues at the edge of the interface are the best targets for affinity optimization, since they possess more mutational flexibility than the amino acids located at the center of the protein-protein interface which often determine the bulk of the affinity of the interaction (37). It has also been shown for antibodies, that in vivo maturation is accompanied by an increased variability of the amino acids of the CDR regions located at the edges of the interface (45). Mutations at these positions exert moderate effects to the association and dissociation kinetics of the interaction (37). The results of our analysis show that the linear
models correlating the logarithm of $k_a$, $k_d$ and $K_d$ to the ZZ scales of the amino acids introduced by mutagenesis, describe and predict the observed data very well. With $Q^2$ values of 0.88, 0.83, and 0.77, they are as good or even better than previously obtained models for kinetic and binding parameters using QSAR methods tested on enzyme/substrate or antibody/oligopeptide systems (26-32;38). The mutants not included in the design (table 3b and 4b (TM, SS, PS & TS(wt)), but within the ZZ scale ranges of the design were predicted successfully within an order of magnitude. This is remarkable since small deviations from the model will lead to a quite large deviation in absolute value, because of the logarithmic transformation of the data. The physical interpretation of the logarithm of the kinetic and affinity constants is an evaluation of the activation and binding energy of the interaction as opposed to the absolute values of $k_a$, $k_d$ and $K_d$

$\Delta G_{k_a}^\circ = -RT\ln k_a$, $\Delta G_{k_d}^\circ = -RT\ln k_d$, $\Delta G^\circ = -RT\ln K_d$ (20;21). In this regard, we consider the model to be a function of energy terms corresponding to the coefficients of the ZZ scales. In energy terms, the standard deviation of the prediction residuals for association, dissociation and affinity is 0.35, 0.49 and 0.56 kcal/mole for all mutants and the four variants not included in the design, indicating that the models have good predictive power. These measurements are surprisingly precise because binding energy measurements have typical errors of 0.5 kcal/mole (24;25) and computational modeling analyses of binding energies of mutants frequently have errors exceeding 3 kcal/mole (9-12;14-16;21), even when training sets are used.

The models quantify clearly the contributions of the ZZ1-3 scales at each position in all three models and no evidence for major interaction between the modified positions was detected. At position 101 small hydrophobic and apolar, non-electrophilic amino acids are optimal for all three parameters of binding and at position 105 different patterns are observed between the three models. A small, hydrophilic (non-electrophilic) amino acid would be optimal for a high $k_a$,
while a polar and electrophilic amino acid would be optimal for a low $k_d$ and $K_d$. This leads to different amino acid preferences for the three models at position 105. Glycine, threonine, glutamine are predicted to be optimal to achieve a high $k_a$. To obtain a low $k_d$, proline, aspartic acid and serine are predicted to be optimal, whereas aspartic acid, serine, glycine and proline would result in an optimal (i.e. low) $K_d$ constant. This clearly shows that the importance of the chemical properties of the amino acid at this position is different when considering activation energy (reflected in the $k_a$ parameter) and the Gibbs free energy of binding ($\Delta G^\circ$) (reflected in the $K_d$ parameter). Energetic contributions of interacting amino acids in protein-protein interactions have mostly been evaluated by the $\Delta G^\circ$ of the interaction. Evaluation of activation energy could also be very useful in elucidating docking trajectories of protein-protein interactions (20;21). Figure 4 reveals that the effect of ZZ1105, important in both the log $k_a$ and log $k_d$ model, is neutral in the log $K_d$ model, due to the positive coefficient (of approximately equal magnitude) of this factor in log $k_a$ and log $k_d$ models ($\log K_d = \log k_d - \log k_a$). Limiting the analysis to the $K_d$ ($\Delta G^\circ$), makes that the importance of this factor cannot be retrieved.

Thr101 in wild-type cAb-Lys3 has a methyl group occupying a relatively hydrophobic cavity in the interface of cAb-Lys3 and lysozyme (fig 2A). In conventional mutagenesis strategies T101S is considered to be a conserved amino acid replacement. We calculated from the crystal structure that this replacement would introduce a cavity of 34 Å$^3$ in the interface. From table 4b we infer that the SS mutant leads to a twenty-fold reduction in affinity compared to TS. The introduction of a cavity results in this case to a decrease of 1.8 kcal/mole, which is about 1 kcal/mole larger than the estimated loss in binding energy for a cavity of this size (24-34 cal/mole Å$^3$) (25). However our QSKR model confirms that ZZ3101 is an important factor in binding. ZZ3 corresponds to the polarity and electrophilicity of the amino acid (27). Serine, as seen from table
1, has a higher ZZ3 value than threonine and therefore sterical aspects do not purely determine
the loss in binding energy of the T101S replacement. This rational is corroborated by the
observation that a valine at position 101 (VN compared to TN in table 4a) leads to a higher
affinity than the wild type residue because of its lower ZZ3 value. This shows that estimation of
the contribution of functional groups by conventional methods, where the functional group is
considered to be independent from the rest of the side chain, would lead to incorrect
interpretation of the contribution to the binding energy of the interaction.

We observed that the wild type amino acid sequence at positions 101 and 105 is quasi optimal
for binding as we can expect from an in vivo maturated antibody. However, our analysis
suggests that the affinity can still be improved. At position 101 a valine (V) is predicted to have a
better affinity than the wild type residue as for position 105, an aspartic acid (D) would be
favored, although predicted differences between wild type TS and the predicted VD mutant are
small (30nM compared to 6nM). It is clear from the comparison of the TN with the VN mutant
(table 4a) that the latter mutant has a twofold lower $k_d$ value than the former. Because of the high
accuracy of $k_d$ constants in surface plasmon resonance (29) we consider this difference to be
significant. This shows that although differences in binding strength are small, QSKR gives
supportive evidence that these small differences in observed affinities and kinetic constants are
true. Extension of this analysis to all interacting residues in the binding interface could result in a
considerable increase in binding strength of the interaction by optimizing a number of contacts
that show small differences in binding energy compared to the wild-type protein. This could also
lead to a rational design strategy to improve affinities or kinetics of binders isolated from naive
or synthetic antibody libraries (1-6). Using a limited set of double or triple mutants and
generating predictive models, we may be able to increase the affinity considerably. We showed
that the combination of two mutated residues in the binding interface can lead to a two thousand fold difference in affinity (fig 3 C, table 4), even though the residues play a secondary role in binding.

Quantitative buffer kinetic relationship (QBKR) & quantitative sequence perturbation relationship (QSPR). In the QBKR analysis we obtain reliable models relating the effect of NaCl, urea, KSCN, DMSO and pH to the kinetics and affinity of the interaction of the cAb-Lys3 variants with lysozyme. The models were converted to chemical sensitivity fingerprints, which were related to the ZZ scale matrix (table 3 a) in a QSPR analysis. Here we probe the effect of ZZ1-3 at positions 101 & 105 in the sequence, to the sensitivity of the binding parameters to the different chemicals. Only the salt sensitivity could be reliably correlated to the ZZ scales. The other models had a low signal to noise ratio as a result of the limited differences in the perturbation of the kinetic or affinity constants by most of the chemical additives.

In the $k_d$ model we see that position 101 does not significantly contribute to the salt sensitivity, while at position 105 a small hydrophilic amino acid increases the salt sensitivity. Smaller amino acids at this position are accompanied by a decrease of the dissociation rate parameter at higher salt levels, whereas larger amino acids are accompanied by an increase in dissociation rate at high salt levels. At this site there are two possible interaction partners for sodium ions at pH above 7.0, namely Asp52, Glu35 of lysozyme (fig. 2 B). In the vicinity of position 101 there are no charged groups present that may interact with sodium ions, possibly leading to a relative insensitivity of the $k_d$ to sodium chloride (fig. 2 A).

In the $k_a$ model for salt sensitivity we see that the larger amino acids at both positions increase the salt sensitivity. The comparison of figures 6 A and 4 A reveals an opposite pattern in the
coefficient plot, showing that mutants leading to a higher salt sensitivity of \( k_a \) also have a lower intrinsic \( k_a \). Since all mutants show an increase of the association rate parameter upon increasing salt concentration, a higher sensitivity of \( k_a \) to salt is accompanied by an increase in \( k_a \). The wild type protein, which shows to have a high association rate and a low NaCl sensitivity, could therefore have optimized the amino acids at these positions in vivo to partially resolve unfavorable long-range electrostatic repulsions in the association process.

Acknowledgements

We thank I. Zegers, D. Maes and M. Lauwereys for reading the manuscript. and M. Malmqvist and M. Hämäläinen for scientific support. M. Vanderveken, M. De Kerpel and C. Bouton for technical assistance. This work was supported by the Vlaams Interuniversitair Instituut voor Biotechnologie and the F.W.O.

References

1. Kurtzman A.L., Govindarajan S., Vahle K., Jones J.T., Heinrichs V. and Patten P.A. (2001) *Curr.Opin.Biotechnol.* **12**, 361-370
2. Dall’Acqua W. and Carter P. (1999) *Current Opinion in Structural Biology* **8**, 443-450
3. Lehmann M. and Wyss M. (2001) *Curr.Opin.Biotechnol.* **12**, 371-375
4. Amstutz P. (2001) *Curr.Opin.Biotechnol.* **12**, 400-405
5. Hanes J., Schaffitzel C., Knappik A., and Pluckthun A. (2000) *Nature Biotechnology* **18**, 1287-1292
6. Griffiths A.D. and Duncan A.R. (1998) *Curr.Opin.Biotechnol.* **9**, 102-108
7. Chowhurdy P.S. and Pastan I. (2002) *Nature Biotechnology* **17**, 568-572
8. Worn A. and Pluckthun A. (2001) *Journal of Molecular Biology* **305**, 989-1010
9. Burnett J.C., Botti P., Abraham D.J., and Kellogg G.E. (2001) *Proteins* **42**, 355-377
10. Novotny J., Bruccoleri R.E., Davis M., and Sharp K.A. (1997) *Journal of Molecular Biology* **268**, 401-411
11. Vajda S., Sippl M., and Novotny J. (1997) *Current Opinion in Structural Biology* **7**, 222-228
12. Jackson T., Morris B.A., Martin A.C., Lewis D.F., and Sanders P.G. (1992) *Protein Engineering* **5**, 343-350
13. Lee L.P. and Tidor B. (2001) *Protein Science* **10**, 362-377
14. Rognan D., Lauemoller S.L., Holm A., Buus S., and Tschinke V. (1999) *Journal of Medicinal Chemistry* **42**, 4650-4658
15. Elcock A.H. (2001) *Journal of Molecular Biology* **312**, 885-896
16. Gromiha M.M., Oobatake M., Kono H., Uedaira H. and Sarai A. (1999) *Protein Engineering* **12**, 549-555
17. Clackson T. and Wells J.A. (1995) *Science* **267**, 383-386
18. Dall'Acqua W., Goldman E.R., Eisenstein E., and Mariuzza R.A. (1996) *Biochemistry* **35**, 9667-9676
19. Dall'Acqua W., Goldman E.R., Lin W., Teng C., Tsuchiya D., Li H., Ysern X., Braden B.C., Li Y., Smith-Gill S.J., and Mariuzza R.A. (1998) *Biochemistry* **37**, 7981-7991
20. Taylor M.G., Rajpal A. and Kirsch J.F. (1998) *Protein Science* **7**, 1857-1867
21. Rajpal A., Taylor M.G., and Kirsch J.F. (1998) *Protein Science* **7**, 1868-1874
22. Ito W., Iba Y., and Kurosawa Y. (1993) *Journal of Biological Chemistry* **268**, 16639-16647
23. Dougan D.A., Malby R.L., Gruen L.C., Kortt A.A., and Hudson P.J. (1998) *Protein Engineering* 11, 65-74

24. Braden B.C., Goldman E.R., Mariuzza R.A., and Poljak R.J. (1998) *Immunol.Rev.* 163, 45-57

25. Braden B.C. and Poljak R.J. (1995) *FASEB J.* 9, 9-16

26. Hansch C. and Leo A. (1995) *Exploring QSAR: Hydrophobic, Electronic, and Steric Constants*, American Chemical Society, Washington DC

27. Sandberg M., Eriksson L., Jonsson J., Sjöström M., and Wold S. (1998) *Journal of Medicinal Chemistry* 41, 2481-2491

28. Hellberg S., Sjöström M., Skagerberg B., and Wold S. (1987) *Journal of Medicinal Chemistry* 30, 1126-1135

29. Andersson K., Choulier L., Hämäläinen M.D., van Regenmortel M.H.V., Altschuh D., and Malmqvist M. (2001) *Journal of Molecular Recognition* 14, 62-71

30. Hansch C. and Klein T.E. (1991) *Methods Enzymology* 202, 512-543

31. Rovero P., Riganelli D., Fruci D., Vigano S., Pegoraro S., Revoltella R., Greco G., Butler R., Clementi S., and Tanigaki N. (1994) *Mol.Immunol.* 31, 549-554

32. Dajani R., Cleasby A., Neu M., Wonacott A.J., Jhoti H., Hood A.M., Modl S., Hersey A., Taskinen J., Cooke R.M., Manchee G.R., and Coughtrie M.W.H. (1999) *Journal of Biological Chemistry* 274, 37862-37868

33. Desmyter A., Transue T.R., Ghahroudi M.A., Thi M.H., Poortmans F., Hamers R., Muyldermans S., and Wyns L. (1996) *Nature Structural Biology* 3, 803-811

34. Decanniere K., Transue T.R., Desmyter A., Maes D., Muyldermans S., and Wyns L. (2001) *Journal of Molecular Biology* 313, 473-478
35. Transue T.R., De Genst E., Ghahroudi M.A., Wyns L., and Muyldermans S. (1998) *Proteins* **32**, 515-522

36. Kabat E.A., Wu T.T., Perry H.M., Gottesman K.S., and Foeler C. (1991) *Sequences of proteins of immunological interest*, US Public Health Services,

37. Lo Conte L., Chothia C., and Janin J. (1999) *Journal of Molecular Biology* **285**, 2177-2198

38. Choulier L., Andersson K., Hämäläinen M.D., van Regenmortel M.H.V., Malmqvist M. and Altschuh D. (2002) *Protein Engineering* (in press)

39. Lauwereys M., Ghahroudi M.A., Desmyter A., Kinne J., Holzer W., De Genst E., Wyns L., and Muyldermans S. (1998) *EMBO J.* **17**, 3512-3520

40. Chen Z. and Ruffner D.E. (1998) *Nucleic Acids Research* **26**, 1126-1127

41. Pace C.N., Vajdos F., Fee L., Grimsley G., and Gray T. (1995) *Protein Science* **4**, 2411-2423

42. Anonymous (1998) *BIAapplications Handbook*, Biacore Ab,

43. Haaland P.D. (1989) *Experimental design in biotechnology*, Marcel Dekker, inc, New York

44. Umetrics AB (1999) *Modde 5.0: User Guide and Tutorial.*, Umetrics AB, Umeå

45. Tomlinson I.M., Walter G., Jones P.T., Dear P.H., Sonnhammer F.E.L., and Winter G. (1996) *Journal of Molecular Biology* **256**, 813-817

Footnotes

1 Abbreviations used are: $k_a$, second order rate constant of association; $k_d$, first order rate constant of dissociation; $K_d$, thermodynamic binding constant; QSAR, quantitative structure
activity relationship; HEL, hen egg-white lysozyme; CDR, complementarity-determining region; cAb, camel single-domain antibody; PBS, phosphate buffered saline; DMSO, dimethylsulfoxide; QSKR, quantitative sequence kinetic relationship; QBKR, quantitative buffer kinetic relationship; QSPR, quantitative sequence perturbation relationship; MLR, multiple linear regression; PLS, partial least squares; ZZ1-3, properties of amino acids (hydrophilicity, size and electronic properties); $R^2$, fraction of the variance described by a model; $Q^2$, predictive power estimate using cross validation.
Figure legends

**Figure 1.** Stereo view of a CPK space filling representation of the lysozyme’s eyview of the cAb-Lys3 binding interface. Residues within 4 Å of lysozyme are color coded: CDR1 residues, pink; CDR2 residues, yellow; CDR3 residues, blue. The side chain atoms of threonine 101 and serine 105 are color-coded: carbon, black; oxygen, red; nitrogen, blue. (from the crystallographically determined coordinates of the cAb-Lys3-Lysozyme complex (Decanniere et al. 2001, pdb code 1JTP (34)).

**Figure 2.** Stereo view of the interactions of threonine 101 (A) and serine 105 (B) of cAb-Lys3 with lysozyme. The lysozyme residues are shown in gray and the cAb-Lys3 residues are colored green. The atoms of threonine 101 (A) and serine 105 (B) are color coded: carbon, black; oxygen, red; nitrogen, blue. Interacting residues within 4 Å of threonine 101 (A) and serine 105 (B) are labeled. (from the crystallographically determined coordinates of the cAb-Lys3-Lysozyme complex (Decanniere et al., pdb code 1JTP (34)).

**Figure 3.** Observed versus predicted values for $k_a$ (A), $k_d$ (B) and $K_d$ (C) and the corresponding residuals versus predicted log $k_a$ (D), log $k_d$ (E) and log $K_d$ (F) of the cAb-Lys3 mutants binding to HEL from the QSKR models. The values of the mutants used to obtain the models are represented by circles (o). The values of the mutants and the wild-type cAb-Lys3 not included in the design are labeled and represented by crosses (+). The ideal correlation is indicated by a straight diagonal line for the observed versus predicted plots (A, B, C) and a horizontal dotted line for the residual plots (D, E, F).
Figure 4. Column plot representation of the MLR coefficients of the QSKR models for log $k_a$ (A), log $k_d$ (B) and log $K_d$ (C). The coefficients were scaled and centered by orthogonal scaling (Modde 5.0: User guide and tutorial (44)). The error bars represent the 95% confidence interval.

Figure 5. Column plot representation of the MLR coefficients of the wild-type cAb-Lys3 models for log $k_a$ (A), log $k_d$ (B) and log $K_d$ (C) of the QBKR analysis, a so called “fingerprint”. The coefficients were scaled and centered by orthogonal scaling to be comparable (Modde 5.0: User guide and tutorial (44)). Both replicate runs were used to obtain the plots. The $R^2$ and $Q^2$ values of the MLR are presented to demonstrate the validity of the models. The error bars represent the 95% confidence interval.

Figure 6. Column plot representation of the PLS coefficients for the NaClSens$k_a$ (A) NaClSens$k_d$ (B) NaClSens$K_d$ (C) models of the QSPR analysis. The coefficients were scaled by unit variance to be comparable (44). The error bars represent the 95% confidence interval.
### Table 1.
The ZZ1-3 scales of the naturally occurring amino acids (from Sandberg et al. (1998)). The last row in the table gives the physico-chemical property represented by the scale.

| Amino acid | ZZ1  | ZZ2  | ZZ3  |
|------------|------|------|------|
| ala        | 0.24 | -2.32| 0.6  |
| arg        | 3.52 | 2.5  | -3.5 |
| asn        | 3.05 | 1.62 | 1.04 |
| asp        | 3.98 | 0.93 | 1.93 |
| cys        | 0.84 | -1.67| 3.71 |
| gln        | 1.75 | 0.5  | -1.44|
| glu        | 3.11 | 0.26 | -0.11|
| gly        | 2.05 | -4.06| 0.36 |
| his        | 2.47 | 1.95 | 0.26 |
| ile        | -3.89| -1.73| -1.71|
| leu        | -4.28| -1.3 | -1.49|
| lys        | 2.29 | 0.89 | -2.49|
| met        | -2.85| -0.22| 0.47 |
| phe        | -4.22| 1.94 | 1.06 |
| pro        | -1.66| 0.27 | 1.84 |
| ser        | 2.39 | -1.07| 1.15 |
| thr        | 0.75 | -2.18| -1.12|
| trp        | -4.36| 3.94 | 0.59 |
| tyr        | -2.54| 2.44 | 0.43 |
| val        | -2.59| -2.64| -1.54|

**Hydrophilicity** | **Size & polarizability** | **Electronic properties:**

Polarity, charge, electrophilicity, electronegativity
Table 2. Buffer composition. In addition to the chemicals listed in the table the buffers also contained 40.6 mM Na₂HPO₄, 9.4 mM KH₂PO₄, 3.34 mM EDTA and 0.005% Surfactant P20 (Biacore AB). Buffers P9, P10 & P11 have identical composition and they constitute a triplicate centerpoint of the design.

| Buffer | NaCl (mM) | KSCN (mM) | Urea (mM) | DMSO (mM) | pH  |
|--------|-----------|-----------|-----------|-----------|-----|
| P1     | 150       | 4         | 40        | 300       | 7.8 |
| P2     | 550       | 4         | 40        | 30        | 7   |
| P3     | 150       | 22        | 40        | 300       | 7.8 |
| P4     | 550       | 22        | 40        | 300       | 7   |
| P5     | 150       | 4         | 400       | 300       | 7   |
| P6     | 550       | 4         | 400       | 30        | 7.8 |
| P7     | 150       | 22        | 400       | 30        | 7   |
| P8     | 550       | 22        | 400       | 300       | 7.8 |
| P9     | 350       | 13        | 220       | 165       | 7.4 |
| P10    | 350       | 13        | 220       | 165       | 7.4 |
| P11    | 350       | 13        | 220       | 165       | 7.4 |
| PBS    | 150       | 0         | 0         | 0         | 7.4 |
Table 3 a. The ZZ scales of the designed mutants

| mutant | ZZ1 (101) | ZZ2 (101) | ZZ3 (101) | ZZ1 (105) | ZZ2 (105) | ZZ3 (105) |
|--------|-----------|-----------|-----------|-----------|-----------|-----------|
| TN     | 0.75      | -2.18     | -1.12     | 3.05      | 1.62      | 1.04      |
| TH     | 0.75      | -2.18     | -1.12     | 2.47      | 1.95      | 0.26      |
| TQ     | 0.75      | -2.18     | -1.12     | 1.75      | 0.5       | -1.44     |
| TA     | 0.75      | -2.18     | -1.12     | 0.24      | -2.32     | 0.6       |
| TG     | 0.75      | -2.18     | -1.12     | 2.05      | -4.06     | 0.36      |
| TP     | 0.75      | -2.18     | -1.12     | 0.27      | 1.84      |           |
| VN     | -2.59     | -2.64     | -1.54     | 3.05      | 1.62      | 1.04      |
| RT     | 3.52      | 2.5       | -3.5      | 0.75      | -2.18     | -1.12     |
| PG     | -1.66     | 0.27      | 1.84      | 2.05      | -4.06     | 0.36      |
| LS     | -4.28     | -1.3      | -1.49     | 2.39      | -1.07     | 1.15      |
| SQ     | 2.39      | -1.07     | 1.15      | 1.75      | 0.5       | -1.44     |
| HD     | 2.47      | 1.95      | 0.26      | 3.98      | 0.93      | 1.93      |
| QP     | 1.75      | 0.5       | -1.44     | -1.66     | 0.27      | 1.84      |
| MV     | -2.85     | -0.22     | 0.47      | -2.59     | -2.64     | -1.54     |

Table 3 b. The ZZ scales of the mutants not included in the design

| mutant | ZZ1 (101) | ZZ2 (101) | ZZ3 (101) | ZZ1 (105) | ZZ2 (105) | ZZ3 (105) |
|--------|-----------|-----------|-----------|-----------|-----------|-----------|
| TS (wild-type) | 0.75 | -2.18 | -1.12 | 2.39 | -1.07 | 1.15 |
| TM     | 0.75      | -2.18     | -1.12     | -2.85     | -0.22     | 0.47      |
| SS     | 2.39      | -1.07     | 1.15      | 2.39      | -1.07     | 1.15      |
| PS     | -1.66     | 0.27      | 1.84      | 2.39      | -1.07     | 1.15      |
Table 4a. Kinetic and affinity constants of the mutants in the design. The measurements were performed in replicate: \( k_a^1, k_a^2 \) represent the \( k_a \) of the first and the second independent measurement, respectively. The same notation was used for the replicates of \( k_d \) and \( K_d \).

|      | \( k_a^1 \) (M\(^{-1}\)s\(^{-1}\)) \(10^4\) | \( k_d^1 \) (s\(^{-1}\)) \(10^3\) | \( K_d^1 \) (nM) | \( k_a^2 \) (M\(^{-1}\)s\(^{-1}\)) \(10^4\) | \( k_d^2 \) (s\(^{-1}\)) \(10^3\) | \( K_d^2 \) (nM) |
|------|---------------------------------|-----------------|----------------|---------------------------------|-----------------|----------------|
| TN   | 3.26                            | 4.14            | 127            | 3.48                            | 4.24            | 122            |
| TH   | 6.97                            | 0.640           | 9.18           | 6.46                            | 0.717           | 11.1           |
| TQ   | 4.09                            | 37.0            | 905            | 3.71                            | 39.2            | 1060           |
| TA   | 4.32                            | 3.74            | 86.6           | 4.68                            | 3.77            | 80.6           |
| TG   | 11.2                            | 1.57            | 14.0           | 12.6                            | 1.53            | 12.1           |
| TP   | 0.375                           | 0.195           | 52.0           | 0.367                           | 0.240           | 65.4           |
| VN   | 3.47                            | 2.04            | 58.8           | 2.99                            | 2.04            | 68.2           |
| RT   | 1.31                            | 25.9            | 1980           | 1.25                            | 24.2            | 1940           |
| PG   | 0.479                           | 21.7            | 4530           | 0.406                           | 21              | 5170           |
| LS   | 3.63                            | 0.528           | 14.5           | 3.68                            | 0.568           | 15.4           |
| SQ   | 1.44                            | 264             | 18300          | 1.46                            | 308             | 21100          |
| HD   | 0.220                           | 3.77            | 1710           | 0.216                           | 3.47            | 1610           |
| TP   | 0.113                           | 0.413           | 365            | 0.0976                          | 0.306           | 314            |
| MV   | 0.759                           | 27.7            | 3650           | 0.774                           | 30.5            | 3940           |

Table 4b. Kinetic and affinity constants of the mutants not included in the design. The measurements were performed in replicate: \( k_a^1, k_a^2 \) represent the \( k_a \) of the first and the second independent measurement, respectively. The same notation was used for the replicates of \( k_d \) and \( K_d \).

|      | \( k_a^1 \) (M\(^{-1}\)s\(^{-1}\)) \(10^4\) | \( k_d^1 \) (s\(^{-1}\)) \(10^3\) | \( K_d^1 \) (nM) | \( k_a^2 \) (M\(^{-1}\)s\(^{-1}\)) \(10^4\) | \( k_d^2 \) (s\(^{-1}\)) \(10^3\) | \( K_d^2 \) (nM) |
|------|---------------------------------|-----------------|----------------|---------------------------------|-----------------|----------------|
| TS (wild-type) | 9.10                            | 0.814           | 8.95           | 7.00                            | 0.889           | 12.7           |
| TM   | 2.32                            | 10.4            | 448            | 2.46                            | 10.7            | 435            |
| SS   | 5.47                            | 9.98            | 182            | 4.34                            | 9.75            | 225            |
| PS   | 0.421                           | 4.80            | 1140           | 0.306                           | 4.76            | 1560           |
Table 5. \( K_s \) Sensitivity fingerprint of the mutants included in the design. The numbers (1) and (2) represent the normalized coefficients of the QBKR models of respectively the first and the second independent replicate.

| Mutant | NaCl    | KSCN    | Urea    | DMSO    | pH      |
|--------|---------|---------|---------|---------|---------|
| TN(1)  | 3.13E-03| -4.65E-04| -1.33E-02| -1.96E-02| -2.01E-02|
| TN(2)  | 3.40E-03| -2.11E-03| -1.33E-02| -1.62E-02| -1.86E-02|
| TH(1)  | 5.02E-03| -2.58E-03| -7.91E-03| -1.93E-02| -1.60E-02|
| TH(2)  | 5.38E-03| 2.69E-04 | -6.88E-03| -1.73E-02| -1.89E-02|
| TQ(1)  | 2.47E-03| 1.95E-04 | -1.30E-02| -2.15E-02| -5.03E-02|
| TQ(2)  | 3.48E-03| -2.67E-03| -1.03E-02| -2.25E-02| -4.44E-02|
| TA(1)  | 8.92E-03| -2.77E-03| -8.79E-03| -1.85E-02| -2.09E-02|
| TA(2)  | 7.86E-03| -6.21E-03| -1.39E-02| -1.80E-02| -2.45E-02|
| TG(1)  | 7.18E-03| -4.11E-03| -9.06E-03| -1.33E-02| -6.51E-03|
| TG(2)  | 8.66E-03| -2.97E-03| -1.16E-02| -1.68E-02| -5.05E-03|
| TP(1)  | 5.52E-03| -4.73E-03| -8.46E-03| -1.46E-02| -1.42E-02|
| TP(2)  | 6.73E-03| -3.38E-03| -5.74E-03| -1.37E-02| -1.64E-02|
| VN(1)  | 2.78E-03| -3.67E-03| -1.56E-02| -1.92E-02| -2.06E-02|
| VN(2)  | 6.71E-03| -5.35E-03| -1.14E-02| -2.02E-02| -2.07E-02|
| RT(1)  | 9.17E-03| -7.75E-04| -1.92E-02| -2.16E-02| -2.64E-02|
| RT(2)  | 1.55E-02| 2.90E-03 | -1.70E-02| -1.55E-02| -2.26E-02|
| PG(1)  | 1.80E-02| -3.15E-03| -1.58E-02| -2.76E-02| -1.15E-02|
| PG(2)  | 1.78E-02| -2.13E-03| -1.79E-02| -2.69E-02| -1.64E-02|
| LS(1)  | 1.50E-02| -5.34E-04| -1.26E-02| -1.86E-02| -1.22E-03|
| LS(2)  | 1.40E-02| -5.86E-04| -1.05E-02| -1.96E-02| -2.44E-03|
| SQ(1)  | 7.91E-03| -6.33E-03| -1.31E-02| -2.65E-02| -6.20E-02|
| SQ(2)  | 8.55E-03| -6.12E-03| -1.42E-02| -2.30E-02| -6.08E-02|
| HD(1)  | 1.56E-02| -1.65E-03| -1.39E-02| -1.38E-02| -6.21E-02|
| HD(2)  | 1.67E-02| -3.53E-03| -1.32E-02| -1.71E-02| -6.74E-02|
| QP(1)  | 2.42E-02| -4.37E-03| -1.24E-02| -1.86E-02| -1.24E-02|
| QP(2)  | 1.21E-02| -1.13E-02| -9.04E-03| -1.22E-02| -1.69E-02|
| MV(1)  | 1.58E-02| -3.84E-03| -1.43E-02| -2.03E-02| -2.05E-02|
| MV(2)  | 1.25E-02| -2.54E-03| -1.57E-02| -2.40E-02| -2.32E-02|
Figures

Figure 1.
Figure 2 A.

Figure 2 B.
Figure 3.
Figure 4.
Figure 5.

A

B

C

NaCl  KSCN  Urea  DMSO  pH

R²=0.948  Q²=0.888

R²=0.960  Q²=0.888

R²=0.963  Q²=0.854
Figure 6.
Kinetic and affinity predictions of a protein-protein interaction using multivariate experimental design
Erwin De Genst, Daphne Areskoug, Klaas Decanniere, Serge Muylldermans and Karl Andersson

J. Biol. Chem. published online June 4, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M202359200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts