

Improving the affinity of antigens for mutated antibodies by use of statistical molecular design

ILONA MANDRIKA, PETERIS PRUSIS, JAKOB BERGSTRÖM, SVIATLANA YAHORAVA and JARL E. S. WIKBERG*

Department of Pharmaceutical Pharmacology, Uppsala University, Box 591, BMC, SE-751 24 Uppsala, Sweden

Received 28 September 2007; Accepted 6 November 2007

Abstract: We demonstrate the use of statistical molecular design (SMD) in the selection of peptide libraries aimed to systematically investigate antigen-antibody binding spaces. Earlier, we derived two novel antibodies by mutating the complementarity-determining region of the anti-p24 (HIV-1) single chain Fv antibody, CB4-1 that had lost their affinity for a p24 epitope-homologous peptide by 8- and 60-fold. The present study was devoted to explore how peptide libraries can be designed under experimental design criteria for effective screening of peptide antigens. Several small peptide-antigen libraries were selected using SMD principles and their activities were evaluated by their binding to SPOT-synthesized peptide membranes and by fluorescence polarization (FP). The approach was able to reveal the most critical residues required for antigen binding, and finally to increase the binding activity by proper modifications of amino acids in the peptide antigen. A model of the active peptide binding pocket formed by the mutated scFv and the antigen was compatible with the information gained from the experimental data. Our results suggest that SMD approaches can be used to explore peptide antigen features essential for their interactions with antibodies. Copyright © 2008 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: single chain antibody; mutagenesis; peptide antigen; SPOT synthesis; fluorescence polarization; statistical molecular design

INTRODUCTION

Owing to their high specificity and ability to recognize virtually any target molecule, antibodies have been found to have many exciting applications, e.g. as biological reagents and therapeutic agents. The interaction of antibodies with their antigens is also a paradigm for the study of mechanisms underlying molecular recognition. The six hypervariable loops or CDRs of the heavy (V_H) and light (V_L) chain variable domains form a unique surface that constitute the antigen binding site that is held in place by a rigid framework of β -sheets [1]. The specificity of an antibody to its antigen relates largely to the conformation of the CDR loops and the identity and positions of their amino acid side chains. A large number of van der Waals forces, hydrophobic interactions, electrostatic forces, and hydrogen bonds take part in the binding of an antigen by the antibody, and these interactions occur over a wide surface area of about 600–900 Å [2].

Different approaches have been developed to obtain selective antigen targets. These include use of synthetic combinatorial libraries [3,4], random peptide libraries,

and phage and ribosome display libraries [5,6]. Such libraries are aimed at generating large numbers of random peptides that are screened to find novel binders. However, it may not be so that larger libraries give a better chance to discover new targets. Several reports have demonstrated that randomly generated libraries with only a limited number of peptides can be useful to discover peptides capable of binding to antibodies with low micromolar or even high nanomolar affinities [7,8]. Moreover, in some cases (e.g. for epitope delineation), small diverse peptide libraries appeared to be superior to large diverse phage libraries [9].

A rational approach to reduce the size of a library is to apply statistical molecular design (SMD). SMD is a technology that aims to reduce the number of entities while preserving the chemical variation in a library. SMD has found use in rational drug design in the selection of sets of compounds for combinatorial and standard organic chemistry [10]. Several authors have applied this approach also to design multiple mutated peptides aimed for the interaction with an antibody [11,12], in the modification of the chemical environment in the characterization of an antibody's antigen binding kinetics [11,13] and in the design of small antibody libraries aimed at the interaction with a native antigen [14]. SMD can be used in combination with multivariate data analysis to provide information on the molecular basis of peptide-antibody interactions and predict the kinetic parameters for the interaction of a peptide and an antibody [15].

The present study was devoted to analyze the utility of SMD in the design of small sets of modified

Abbreviations: FP, fluorescence polarization; scFv, single chain fragment variable; CDR, complementarity determining region; CB4-1, murine monoclonal anti p-24 (HIV-1) antibody; h-peptide, epitope-homologous peptide; V_L , light chain variable fragment; PyBOP, benzo triazol-1-yl-oxy-tris-pyrrolidino-phosphonium-hexafluorophosphate.

* Correspondence to: Jarl E. S. Wikberg, Department of Pharmaceutical Pharmacology, Uppsala University, Box 591, BMC, SE-751 24 Uppsala, Sweden; e-mail: Jarl.Wikberg@farmbio.uu.se

peptides aimed to investigate the importance of peptide antigen amino acids for interaction with wild-type (wt) and mutated antibodies. As a study case, we used two single chain antibodies from the CB4-1 scFvs library described in one of our preceding studies [16]. These antibodies had been derived by mutations from the CB4-1 antibody directed to a HIV-capsid protein derived peptide termed h-peptide [17]. The two selected mutants had lost their affinity by 8- and 60-fold for the h-peptide and the present study was aimed to devise a strategy to increase activity of the peptide target for these antibodies by applying SMD.

MATERIALS AND METHODS

Antibody Mutants

In a previous study, we had created a library of multiple mutated scFvs based on the wt scFv-CB4-1 [16]. In this library, four residues of the CDR loop 3 of the light chain being in close proximity with h-peptide residues had been simultaneously mutated. From this library, we selected two mutated antibodies (here termed NQL and EEP antibodies), which showed dissociation constants of about 0.2 and 1.2 μM for the wt h-peptide antigen. This corresponded to an 8- and 60-fold loss of affinity compared to the wt antibody. Both the selected antibodies had mutations at positions 92–94 in the CDR L3 loop. For the NQL antibody, Asp⁹² was mutated to Asn, Asp⁹³ to Gln and Phe⁹⁴ to Leu, and for the EEP antibody both Asp⁹² and Asp⁹³ were mutated to Glu, while Phe⁹⁴ was mutated to Pro.

Statistical Molecular Design of h-Peptide based Peptide Libraries

We created five sets of peptides using SMD. SMD is a sampling technique for molecular entities, which can select subsets of entities from a full set by applying experimental design techniques in order to represent the full set as good as possible [18,19].

The starting point for the SMD was the description of the properties of the peptide amino acids by the five amino acid z-scales by Sandberg *et al.* [20]. These z-scales are principal components obtained by compression of 26 measured and computed physicochemical and structural properties of amino acids by principal component analysis [20]. The Sandberg z-scales represent approximately hydrophilicity (z1), steric/bulk properties and polarizability (z2), polarity (z3), and electronic properties (z4, z5) [20].

Since the five scales represent orthogonal vectors it might be enough to explore three points along each scale – two points at the extremes and one in the middle. Thus, the design is reduced to having one representative amino acid for each extreme value and one for the middle value. Such a selection could be the best choice if one assumes that the binding activities are linearly represented at the extremes. However, this is a highly unlikely situation as the binding surface for peptide–antibody interactions should likely be very nonlinear. Then, a more safe approach would be to explore amino acids that reside between extremes and middle values. In order to find out set of such amino acids, we considered amino acids having z-scale value close to the middle values. If one selected different amino acids for each z-scale one would need 15 different amino acids (i.e. 3 amino acid times 5 z-scales), which would result in a huge library of peptides although certainly smaller than selecting all possible combinations of 20 amino acids per position. In order to overcome this problem, we decided to let the amino acid chosen to represent several z-scales. Another concern in the selection of amino acids was to avoid using His, as an anti-His secondary antibody was used in the immunoassays of the His-tagged scFvs, and peptides with His would, therefore, interfere with the assays.

Taking all these considerations into account, we came up with two general sets of amino acids. The first set was designed to be as small as possible and contained only six amino acids, namely, Ala, Lys, Met, Ser, Tyr, and Val (design set 1; Table 1). In the second set, we increased the coverage of the chemical space of amino acids by representing the extremes of the z-scales with different amino acid. This yielded a set of ten amino acids, namely, Ala, Asp, Lys, Glu, Gln, Met, Phe, Ser, Trp, and Val (design set 2; Table 2).

Table 1 Six amino acid design set with their z-scales

Value range	z1	z2	z3	z4	z5
Upper (+)	SER (2.39)	TYR (2.44)	SER (1.15)	MET (1.94)	ALA (1.30)
Middle (0)	ALA (0.24)	MET (−0.22)	TYR (0.43)	TYR (0.04)	VAL (−0.02)
Lower (−)	MET (−2.85)	VAL (−2.64)	LYS (−2.49)	SER (−1.39)	TYR (−1.47)

Table 2 Ten amino acid design set with their z-scales

Value range	z1	z2	z3	z4	z5
Upper (+)	SER (2.39)	TRP (3.94)	ASP (1.93)	TRP (3.44)	ALA (1.30)
Middle (0)	ALA (0.24)	MET (−0.22)	GLU (−0.11)	PHE (0.54)	GLU (−0.25)
Lower (−)	MET (−2.85)	VAL (−2.64)	LYS (−2.49)	GLN (−1.34)	TRP (−1.47)

Initially, we explored three peptide libraries by altering amino acids at positions 3–5 (*Lib 1*), at positions 6–8 (*Lib 2*) and at positions 9–11 (*Lib 3*) of the h-peptide, using design set 1 of 6 amino acids (see Table 1). However, as the original amino acids did not necessarily occur in our design set, we replaced some of the amino acids from the design set so that the original amino acids were included. This gave slightly modified designs for each position, depending on the original amino acids of that position (Table 3). The result obtained from the testing of *Lib 1–3* prompted the need for exploration of positions 9 and 10, using a more diverse set of amino acids for which we accordingly elected to use design set 2. However, as some of the amino acid combinations proved synthetically difficult, we made slight changes in original design set to improve the synthetic yield. Thus, for position 9 we elected to use asparagine rather than aspartic acid, as well as we choose proline rather than methionine at this position. For position 10, we elected to use isoleucine rather than valine. The final design sets for positions 9 and 10, which were used in the design of *Lib 4* are listed in Table 3.

Following the evaluation of *Lib 4* we selected three modified peptides, Gln⁹Ala¹⁰, Pro⁹Ser¹⁰, and Asn⁹Ile¹⁰, for further modifications at positions 3–5 using design set 1, which resulted in *Lib 5*.

The automatic synthesizer used in the present study suggested using libraries with multiples of 24 peptides. For the *Lib 2* we synthesized 24 peptides, for *Lib 1*, *Lib 3*, and *Lib 4* 48 peptides were synthesized while for *Lib 5* we synthesized 3 × 48 = 144 peptides. D-optimal design [21] was applied on the selected amino acids using the respective design sets, which resulted in the respective library. The final libraries are listed in Tables 4 and 5 and Figure 3(a).

Table 3 Amino acids selected for multiple substitutions of h-peptide residues 3–11 according to statistical molecular design

Amino acid position in h-peptide	Amino acids selected for substitution	
Thr ³	Ala, Lys, Met, Ser, Tyr, Val ^a	<i>Lib 1</i>
Pro ⁴	Lys, Met, Ser, Tyr, Val, Ala ^a	
Glu ⁵	Ala, Lys, Met, Tyr, Val, Ser ^a	
Asp ⁶	Ala, Lys, Met, Tyr, Val, Ser ^a	<i>Lib 2</i>
Leu ⁷	Ala, Met, Ser, Tyr, Val, Lys ^a	
Asn ⁸	Ala, Lys, Met, Tyr, Val, Ser ^a	
Gln ⁹	Ala, Met, Ser, Tyr, Val, Lys ^a	<i>Lib 3</i>
Lys ¹⁰	Ala, Met, Ser, Tyr, Val, Lys	
Leu ¹¹	Ala, Met, Ser, Tyr, Val, Lys ^a	
Gln ⁹	Ala, Glu, Phe, Leu, Asn, Pro, Ser, Val, Trp	<i>Lib 4</i>
Lys ¹⁰	Ala, Glu, Asp, Ile, Lys, Gln, Met, Ser, Trp	

^a Amino acid removed from design set 1 and replaced by the original h-peptide amino acid. The replacement was done by comparing z1–z5 scales of the original amino acid with amino acids selected for substitutions. The amino acid representing similar z scale directions (low, medium or high) as the original one was removed from the design set.

Expression of CB4-1 Single Chain Mutants

The scFv CB4-1 mutants were cloned into the expression vector pET 22b(+) (Novagen). This vector carries the sequence of a signal peptide PelB that permits periplasmic expression of the scFv CB4-1 in *E. coli*, and a His-tag for detection and purification purposes.

ScFvs were grown in the BL21 strain of *E. coli* in 2 × YT medium containing 100 µg/ml ampicillin and 1% glucose. After reaching OD₆₀₀ 0.5–0.6, antibody expression was induced by the addition of 0.05 mM IPTG, followed by overnight incubation at room temperature. (Prior to induction, the cells were washed to remove glucose.) The soluble fractions of the scFvs were obtained from the periplasms by osmotic shock according to the pET vector manual (Novagen). The fractions were dialyzed overnight against 30 volumes Na₂HPO₄ buffer, pH 8.0 at 4 °C and purified on Ni-NTA agarose gel (Qiagen). The eluted scFvs were dialyzed overnight against 50 volumes of 50 mM Tris-HCl, pH 8.0; 5 mM EDTA and further purified on a Mono-Q FPLC column (Amersham Biosciences). The scFv fractions were concentrated to 0.2–0.5 mg/ml on 10 kDa Vivaspin concentrators (Vivascience AG). Antibody purity was checked by SDS-PAGE. Protein concentrations were determined using the BCA protein assay kit (Pierce).

SPOT Synthesis of Peptide Sets on Cellulose Membranes

SPOT synthesis was carried out using an automated multiple peptide synthesizer (MultiPep, Intavis AG Bioanalytical Instruments, Germany) using Amino-PEG₅₀₀ cellulose membranes (Intavis AG Bioanalytical Instruments, Germany). The peptides were synthesized using N^α-Fmoc amino acids. The side chain protecting groups were the following: Trt (Asn, Gln), tBu ether (Ser, Thr, Tyr), OtBu ester (Asp and Glu), Boc (Lys, Trp). Fmoc amino acids were pre-activated daily by incubating 0.15 mmol of the amino acid with 330 µl 0.75 M HOBt followed by dilution to 450 µl with 150 µl of 1.1 M DIPCl. All solutions were in N-methyl-2-pyrrolidone. Coupling efficiency was monitored after each cycle by bromophenol blue staining (0.005% solution in DMF). After the peptide sequences had been assembled, the side chain protecting groups were removed by treatment with deprotection mixture (TFA – DCM – triisopropylsilane – water, 7.5:7.5:0.45:0.3 ml). This treatment was followed by four washes with DCM, four with DMF and two with EtOH (2 min). The membrane was then dried and kept at –20 °C before use.

Immunoassay with Cellulose-Bound Peptides

The membrane arrays were rinsed with ethanol for 1 min and washed 3 times with Tris-buffered saline (TBS, 50 mM Tris-HCl, pH 8.0, 137 mM NaCl, 2.7 mM KCl) for 10 min. The membranes were then blocked overnight with TBS containing 0.05% Tween 20 (T-TBS), 3% dry milk, 1% BSA, and 1% sucrose. After washing with T-TBS, 0.5 µg/ml of scFvs were added in blocking buffer and incubated for 3 h at RT. The membranes were then washed 3 times with T-TBS, and peroxidase-labeled anti-rabbit-His antibody (Sigma, Sweden) was applied at 1:9000 dilution for 2 h at RT. After washing with T-TBS the membranes were incubated with ECL chemiluminescence substrate (Amersham Biosciences,

Table 4 Peptide sequences of *Lib 1–3*

GAXXXDLNQLAGN (<i>Lib 1</i>)		GATPEXXXQKLAGN (<i>Lib 2</i>)		GATPEDLNXXXAGN (<i>Lib 3</i>)	
1. TPE ^a	25. MMM	1. DLN ^a	25. MMM	1. QKL ^a	25. MMM
2. SYM	26. SSA	2. DSM	26. SSA	2. SYM	26. SSA
3. TYK	27. ASV	3. DLK	27. ASV	3. QYS	27. ASV
4. MSK	28. YPA	4. VLV	28. YPA	4. MSS	28. YKA
5. YMM	29. SYA	5. ALN	29. SYA	5. YMM	29. SYA
6. TSM	30. KYV	6. MLV	30. KYV	6. QSM	30. VYV
7. TPK	31. MKY	7. AAV	31. MKY	7. QKS	31. MAY
8. SPV	32. MME	8. KLK	32. MME	8. SKV	32. MML
9. KKA	33. MYA	9. VMK	33. MYA	9. VAA	33. MYA
10. YMY	34. YYE	10. KYM	34. YYE	10. YMY	34. YYL
11. TMV	35. YKK	11. YVM	35. YKK	11. QMV	35. YAS
12. TVA	36. AVK	12. DLY	36. AVK	12. QVA	36. AVS
13. APE	37. MVY	13. VYA	37. MVY	13. AKL	37. MVY
14. MPV	38. AYY	14. KYV	38. AYY	14. MKV	38. AYY
15. AKV	39. SVE	15. MAY	39. SVE	15. AAV	39. SVL
16. KPK	40. APM	16. MMN	40. APM	16. VKS	40. AKM
17. YSE	41. AKK	17. MVY	41. AKK	17. YSL	41. AAS
18. KSM	42. KVE	18. AYY	42. KVE	18. VSM	42. VVL
19. KMY	43. AMA	19. VVN	43. AMA	19. VMY	43. AMA
20. TPY	44. TKV	20. ALK	44. TKV	20. QKY	44. QAV
21. SMK	45. YVV	21. YVV	45. YVV	21. SMS	45. YVV
22. KYM	46. SKM	22. VAM	46. SKM	22. VYM	46. SAM
23. YVM	47. KSA	23. KSA	47. KSA	23. YVM	47. VSA
24. SSY	48. TKE	24. DAN	48. TKE	24. SSY	48. QAL

^a sequence of the wt h-peptide.**Table 5** Peptide sequences of the *Lib 5* containing 3 sets of 48 peptides each. The 3 modified h-peptides as indicated in the table's top line were further substituted at positions 3–5 as indicated in the subsequent lines

GAXXXEDLNQALAGN		GAXXXEDLNPSLAGN		GAXXXEDLNNILAGN	
1. TPE	25. TMA	49. TPE	73. MMA	97. TPE	121. MME
2. YPE	26. MSA	50. TKE	74. SSA	98. SPE	122. YSE
3. SKE	27. YSA	51. SME	75. AYA	99. AKE	123. AYE
4. YKE	28. KYA	52. YME	76. KVA	100. KME	124. TVE
5. YVK	29. KSY	53. SPM	77. AMY	101. KPM	125. KYY
6. APM	30. SSY	54. YPM	78. TSY	102. MKM	126. AVY
7. YKM	31. AYY	55. KKM	79. MYY	103. AMM	127. SVY
8. KMM	32. TVY	56. MMM	80. YYY	104. SSM	128. MPV
9. AME	33. AVA	57. ASE	81. YVA	105. MPA	129. SYA
10. MSE	34. KPK	58. KSE	82. APK	106. AKA	130. YYA
11. TYE	35. SPK	59. KYE	83. YKK	107. YMA	131. TVA
12. KVE	36. AKK	60. SYE	84. TMK	108. KSA	132. MVA
13. TSM	37. TPV	61. SMM	85. YVY	109. TYM	133. MKV
14. MYM	38. KPV	62. ASM	86. YPV	110. YVM	134. SKV
15. MVM	39. KKV	63. YSM	87. AKV	111. TPY	135. TMV
16. SVM	40. SMV	64. AYM	88. KMV	112. APY	136. AMV
17. MPE	41. MKK	65. AVE	89. SSK	113. YPK	137. MSK
18. KPA	42. MMK	66. TPA	90. TYK	114. TKK	138. MYK
19. SPA	43. TSK	67. MKA	91. KYK	115. SMK	139. KVK
20. TKA	44. SYK	68. AMA	92. MVK	116. ASK	140. SVK
21. MPY	45. ASV	69. TVM	93. MSV	117. KKY	141. TSV
22. TKY	46. MYV	70. KVM	94. TYV	118. YKY	142. YSV
23. KMY	47. YVV	71. KPY	95. AVV	119. MMY	143. SYV
24. YMY	48. YVV	72. SKY	96. SVV	120. MSY	144. KVV

Sweden) where after spots were visualized using a CCD-camera.

Standard Solid-Phase Peptide Synthesis

Soluble peptides were synthesized in 5 μmol scale on a MultiPep synthesizer (Intavis AG Bioanalytical Instruments, Germany) using its automated standard protocol optimised for Fmoc chemistry. PyBOP was used as activating reagent and Rink Amide MBHA resin (Calbiochem-Novabiochem AG, Switzerland) as polymeric support. All peptides were characterized by analytical HPLC on a Waters system (Waters, Milford, MA, USA) equipped with a Vydac C18 reversed-phased column. Purity of raw peptides was according to HPLC above 80%. Their structures were confirmed by LC/MS spectrometer (PerkinElmer Life and Analytical Sciences, Boston, MA, USA). All peptides were prepared as their C-terminal amides.

For the ELISA technique, the peptides were biotinylated at the amino terminus via a spacer arm (ϵ -aminocaproic acid) using the standard protocol for Fmoc chemistry.

For FP assay, the peptides were labeled at the amino terminus with fluorescein succinimidyl ester (Pierce) and purified by reversed-phase HPLC.

ELISA Assay

Streptavidine-coated plates (Thermolabsystem) were incubated overnight at 4 °C with 100 μl per well of a 5 $\mu\text{g}/\text{ml}$ of biotinylated substituted h-peptides in 1 \times phosphate-buffered saline (PBS), pH 7.4. After the incubation, the plates were washed 5 times with PBS/0.05% Tween-20, and the wells were then incubated for 1 h with PBS/1% BSA, in order to block nonspecific antibody binding. The scFvs (1 $\mu\text{g}/\text{ml}$, diluted in PBS/1% BSA) were then added and incubation proceeded for 2 h at room temperature. After washing, goat anti-rabbit-His-HRP antibody (Sigma), diluted 1:9000 in PBS/1% BSA, was added and incubated for 1 h. After washing, the bound antibodies were detected by adding HRP substrate (i.e. *o*-phenylenediamine/ H_2O_2 in citrate-phosphate buffer, pH 5.0) and incubating for 10 min. The reaction was stopped by adding 30 μl of 2 M H_2SO_4 and absorbance was read at 495 nm using a Vmax kinetic microplate reader (Novo BioLabs).

Fluorescence Polarization Binding Assays

All assays were performed in black 96-well plates (Costar). Mixtures containing a fixed concentration of labeled peptide and different concentrations of scFvs (ranging 1.4–356 nM) were incubated at room temperature for 10 min in PBS, pH 7.4, containing 0.05% Tween 20 and 0.1% BSA. (Further incubation up to 2 h did not change the signal intensity). FP measurements were performed on a POLARStar Optima counter (BMG), using a 485 nm excitation filter and a 520 nm emission filter. Dissociation constants were obtained from the binding curves by nonlinear least square curve fitting using a one site binding equation in GraFit 5.0 software [22].

Mutant scFv–Mutant Peptide Complex Modeling

Structure models for the antibody–peptide binding complexes were generated by homology modeling using the X-ray structure of the CB4-1 Fab in complex with the h-peptide

as a template (PDB entry code 1BOG) [23]. The substitution of the mutated amino acid residues in both peptides and antibody were performed using the O-program [24]. The models were then submitted to a server-based program NOMAD-Ref for steepest descent energy minimization using the Gromos96 force field [25]. Finally, the structures of scFv–peptide complexes were viewed using Pymol [26]. The antibodies amino acid numbering scheme used was according to Kabat *et al.* [27].

RESULTS

Effect of scFv Mutations on h-Peptide Binding

Mutating the two antigen contacting residues (Asp⁹² and Phe⁹⁴) and the noncontacting residue (Asp⁹³) of the CDR L3 loop of the CB4-1 antibody binding resulted in reduced affinity for h-peptide as revealed by FP assays. While the wt CB4-1 antibody bound to the h-peptide with a K_d of 21 nM, the NQL and EEP mutants bound it with K_d s of 160 and 1210 nM, respectively (Table 6). Thus, removing the negative charges at positions 92 and 93, and replacing the phenylalanine at position 94, which forms an aromatic interaction with Lys¹⁰ of the h-peptide, with leucine in the CDR L3 loop (i.e. NQL antibody) resulted in an 8-fold reduction in the affinity of the h-peptide. Keeping the negative charges at positions 92 and 93, but replacing the Phe⁹⁴ with proline (i.e. EEP antibody) reduced the affinity 60-fold. These results indicate that hydrophobic or aromatic side chains are important at position 94 for the binding of the h-peptide, and that the aromatic function is most important.

Effects of Multiple Amino Acid Substitutions in the h-peptide on scFvs Binding

We first made multiple changes in the middle part of the h-peptide to elucidate the importance of this region for the binding to wt and mutant scFv antibodies. This was done by substituting residues Asp⁶, Leu⁷, and Asn⁸ of the h-peptide with the amino acids shown in Table 3, making the selections with D-optimal design. This resulted in *Lib 2* of 24 peptides (Table 4). The 24 peptides were synthesized as spots on cellulose membranes and the ability of our three antibodies to bind to them was evaluated. The results of this analysis showed that replacement of Leu⁷ with any other amino acid selected from the design set completely abolished antibody binding. The only peptides that showed some binding activity had the original Leu⁷ in their sequences. Moreover, the results from this analysis showed that only the wt antibody, but not the mutated ones, bound to peptides with single amino acid substitutions (Asn⁸Tyr, Asn⁸Lys, and Asp⁶Ala). These results suggest that Leu⁷ influences the binding affinity of the h-peptide critically for both wt and mutated

antibodies as well as Asp⁶ or Asn⁸ residues play a critical role for the h-peptide binding, in particular for both of the mutated scFvs.

In the next experiments, we analyzed the contribution of the *N*- and *C*-terminal parts of the h-peptide to the antibody binding. For the *N*-terminal part Thr³, Pro⁴, and Glu⁵ were selected while for the *C*-terminal part Gln⁹, Lys¹⁰, and Leu¹¹ were selected for substitutions applying the same approach as for designing the substitutions for *Lib 2*, except that the D-optimal design covered 48 amino acids for both sets. The resulting libraries, *Lib 1* and *Lib 3* are shown in Table 4. These were synthesized on cellulose membranes using the SPOT technology and evaluated for antibody binding (Figure 1). Several highly and fairly reactive peptides identified on the SPOT membranes were re-synthesized and their ability to interact with the antibodies was confirmed by ELISA (Figure 2). As seen from Figures 1(a) and 2(a) almost all multiple modifications at the *N*-terminal part led to peptides that retained, or only slightly lost, their abilities to bind to the wt scFv. However, the mutated antibodies proved to be very sensitive to the *N*-terminal substitutions as they bound slightly only to 2 peptides of this type, bearing single amino acid substitutions in their *N*-terminus, namely, Thr³Ala and Glu⁵Tyr (p13 and p20, *Lib 1*) (Figures 1(a) and 2(a)).

For the *C*-terminally substituted h-peptides only a few binders were obtained for both wt and mutated antibodies (Figures 1(b) and 2(b)). The best binding peptide had the single amino acid substitution Lys¹⁰Ala (p48, *Lib 3*) and it retained binding to the wt antibody, albeit with a decrease in binding efficiency (Figure 1(b)). Moreover, the NQL mutant bound equally well to

Lys¹⁰Ala (p48, *Lib 3*) as it did to the original h-peptide, while the EEP mutant showed somewhat better binding to the Lys¹⁰Ala peptide compared to the h-peptide (Figure 2(b)). None of the other peptides from *Lib 1* or *3* exhibited any binding activity over the noise level for the mutant scFvs, whether assessed by ELISA or SPOT.

Influence of Gln⁹Lys¹⁰ Substitutions on scFvs Binding

As the substitutions of the h-peptide at position Lys¹⁰ gave peptides with about the same or even slightly enhanced binding to the mutated antibodies as the h-peptide, we decided to explore this region further by creating a library of single Lys¹⁰ and double Gln⁹, Lys¹⁰ substitutions.

To this end, we first synthesized a small library comprising 48 Gln⁹Lys¹⁰ mutants (*Lib 4*) on SPOT membranes (Figure 3(a)). All of our three antibodies showed binding activities to some of the peptides of this library, although most binders were obtained for the wt and NQL antibodies (Figure 3(b)). In order to verify the results obtained on the SPOT membranes, seven peptides were selected for conventional solid-phase peptide synthesis and used to assess the binding activities by FP (Table 6). As seen from these exhibited measurable binding, the K_d ranges around 10^{-8} – 10^{-6} M (Table 6). Moreover, as expected, peptides that did not show binding activity for the EEP antibody on SPOT membrane failed to bind in the FP assay as well (*cf* Figure 3 and Table 6). However, one peptide that reacted with the NQL antibody in the SPOT assay, namely, Val⁹Asp¹⁰ (p43, *Lib 4*) did not show

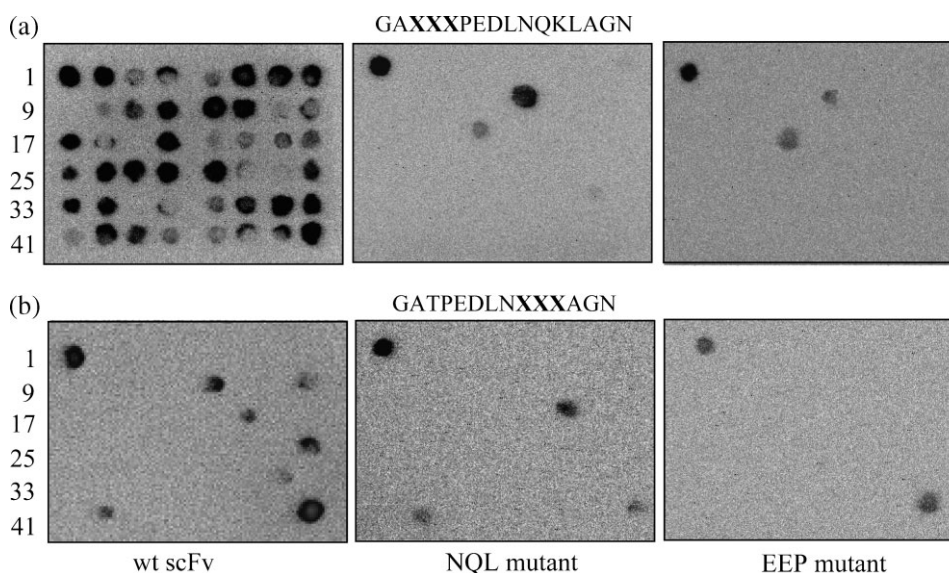


Figure 1 Binding of wt and mutated scFv antibodies to cellulose bound peptides. Peptides were modified at the *N*- and *C*-terminals of the h-peptide by substitutions at positions 3–5 (*Lib 1*) (a) and 9–11 (*Lib 3*) (b). Each membrane represents a set of 48 peptides, where each peptide had been synthesized as an individual spot. The antibody binding was detected with second antibody based chemiluminescence assay and visualized using a CCD-camera.

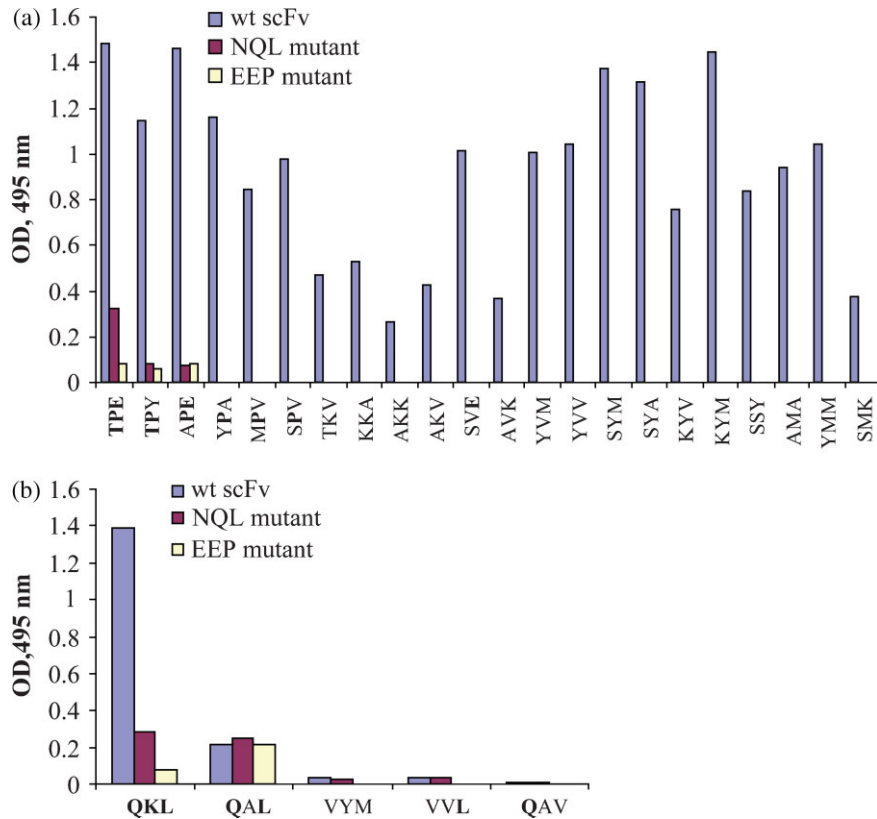


Figure 2 Influence of multiple h-peptide substitutions on wt and mutated scFv antibody binding measured by ELISA. Shown are the binding activity of h-peptides N-terminally substituted at positions 3, 4, and 5 (a) and C-terminally substituted at positions 9, 10, and 11 (b). The h-peptide wt sequence is marked in bold face. Binding of the scFvs carrying the His-tag was detected using a peroxidase-conjugated anti-His antibody.

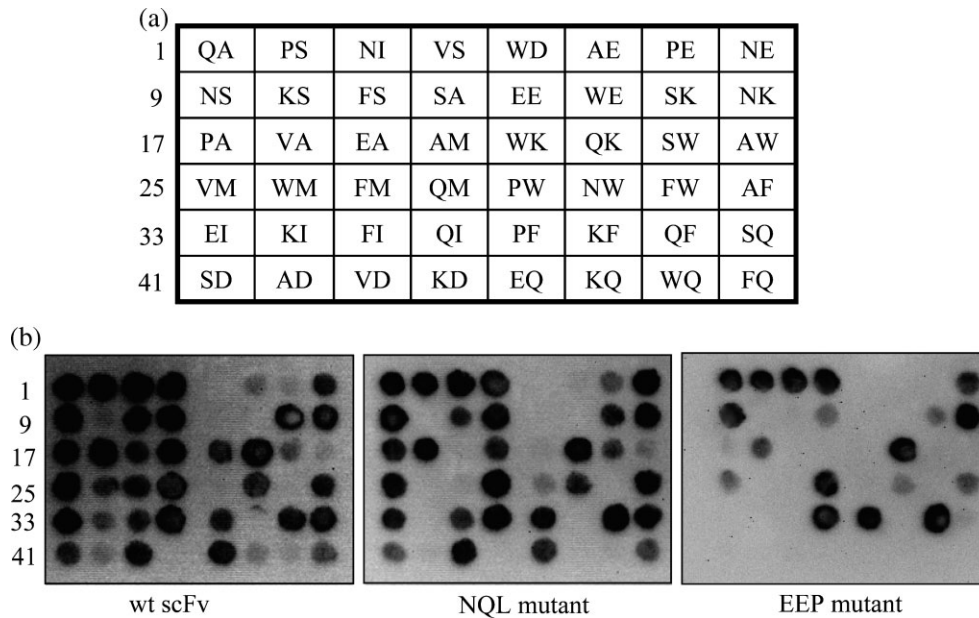


Figure 3 Binding of the wt and mutated scFv to cellulose bound peptides derived from the h-peptide sequence (*Lib 4*). The sequences of the 48 h-peptide substitutions at positions 9 and 10 are indicated in panel (a). Each peptide was synthesized as an individual spot on cellulose membranes and assayed for the binding of the wt and mutated antibodies (b). Antibody binding was detected with second antibody based chemiluminescence assay and visualized using a CCD-camera.

Table 6 Binding affinity constants of substituted h-peptides for wt and mutant scFvs

Peptide sequence	Binding to wt scFv K_d , nM ^a	Binding to NQL mutant, K_d , nM	Binding to EEP mutant K_d , nM	
GATPEDLN Q KL A GN	21 ± 5	160 ± 28	1210 ± 132	wt
GATPEDLN Q A L A GN	261 ± 8	153 ± 10	364 ± 20	p1, <i>Lib 4</i>
GATPEDLN P S L A GN	312 ± 14	107 ± 18	757 ± 29	p2, <i>Lib 4</i>
GATPEDLN N I L AGN	60 ± 18	98 ± 25	1211 ± 328	p3, <i>Lib 4</i>
GATPEDLN N S L A GN	85 ± 19	165 ± 17	902 ± 96	p9, <i>Lib 4</i>
GATPEDLN A M L AGN	212 ± 25	515 ± 88	NM	p20, <i>Lib 4</i>
GATPEDLN Q F L A GN	220 ± 3	145 ± 15	133 ± 3	p39, <i>Lib 4</i>
GATPEDLN V D L AGN	1836 ± 97	NM	NM	p43, <i>Lib 4</i>
G A S K E D L N Q A L A GN	249 ± 35	872 ± 69	NM	p3, <i>Lib 5</i>
G A A P M D L N Q A L A GN	367 ± 26	1260 ± 359	NM	p6, <i>Lib 5</i>
G A S P A D L N Q A L A GN	238 ± 83	555 ± 154	1695 ± 201	p19, <i>Lib 5</i>
G A S P M D L N P S L A GN	866 ± 104	380 ± 89	589 ± 75	p53, <i>Lib 5</i>
G A S K Y D L N P S L A GN	1100 ± 250	403 ± 56	NM	p72, <i>Lib 5</i>
G A K V A D L N P S L A GN	924 ± 123	512 ± 73	NM	p76, <i>Lib 5</i>
G A A P Y D L N N I L AGN	405 ± 70	1509 ± 436	NM	p112, <i>Lib 5</i>

Positions substituted in the h-peptide sequence are shown in bold face.

^a Binding dissociation constants, K_d , were determined by fluorescence polarization. The data shown are the mean ± SD calculated from two or three independent experiments each performed in duplicate.

NM, Not measurable

any detectable binding in the fluorescence polarization assay (Figure 3; Table 6). These results are compatible with the earlier suggestion that antibody binding is facilitated by the high density of peptide at the surface of cellulose membrane [28], thus making this type of assay highly sensitive.

Another notable peptide was the doubly substituted Gln⁹Pro/Lys¹⁰Ser peptide (p2, *Lib 4*), as it showed almost 3-fold selectivity for the NQL mutant compared to the wt antibody (Table 6). Yet, another interesting one was the Lys¹⁰Phe substitution (p39, *Lib 4*) which showed an almost 10-fold increase in the affinity for the EEP mutant, while for the wt antibody the affinity decreased about 10-fold compared to the h-peptide (Table 6).

Of the several substituted peptides that showed improved affinity and/or improved selectivity for the mutated antibodies we selected three C-terminal mutated peptides, namely, Lys¹⁰Ala (p1, *Lib 4*), Gln⁹Pro/Lys¹⁰Ser (p2, *Lib 4*) and Gln⁹Asn/Lys¹⁰Ile (*Lib 4*, p1–3), for further substitution analysis at positions 3–5 (*Lib 5*; Table 5). *Lib 5* was first SPOT synthesized; assay results for the scFv antibodies are shown in Figure 4. Several of the active peptides were re-synthesized for fluorescence polarization assaying (Table 6). Despite that positions 3–5 were modified according to the same design principles for both *Lib 1* and *Lib 5*, *Lib 5* showed a much larger number of binders for mutated antibodies than *Lib 1* (Figures 4 and 1(a)). For the NQL antibody, almost the entire library of N-terminal substituted double

mutants Gln⁹Pro/Lys¹⁰Ser and Gln⁹Asn/Lys¹⁰Ile was reactive in the SPOT analysis. However, the EEP antibody bound only to multiply substituted peptides where Pro⁴ remained intact and Thr³ was conservatively substituted by the physicochemically similar Ser (Figure 4). A few binders selected from *Lib 5* showed binding affinity in the 10⁻⁶–10⁻⁷ M range in the FP assay (Table 4). Accordingly, it seems that the peptides substituted at positions 9–10 tolerate multiple N-terminal substitutions better than does the h-peptide.

Possible Explanation of the Mutational Effects by Structure

The roles of the multiple amino acid substitutions in the antibody–peptide binding site were further investigated by modeling the antibody–peptide structure. In the wt antibody/h-peptide complex (Figure 5(a)), Lys¹⁰ stacks in a hydrophobic environment formed by surrounding residues and makes a hydrophobic interaction via its methylene group with aromatic side chain of the CDR L3 loop Phe⁹⁴. The replacement of the Phe⁹⁴ aromatic side chain by the smaller hydrophobic Leu residue in the NQL mutant or Pro in the EEP mutants could presumably partially distract the hydrophobic cavity, which could result in reduction of hydrophobic interactions (loss of some van der Waals contacts) between the Leu⁹⁴ or Pro⁹⁴ antibody residues and the Lys¹⁰ peptide residue. As a result, the affinity of both mutants for wt h-peptide is reduced. Substitutions at position 10 of the h-peptide with amino acids with

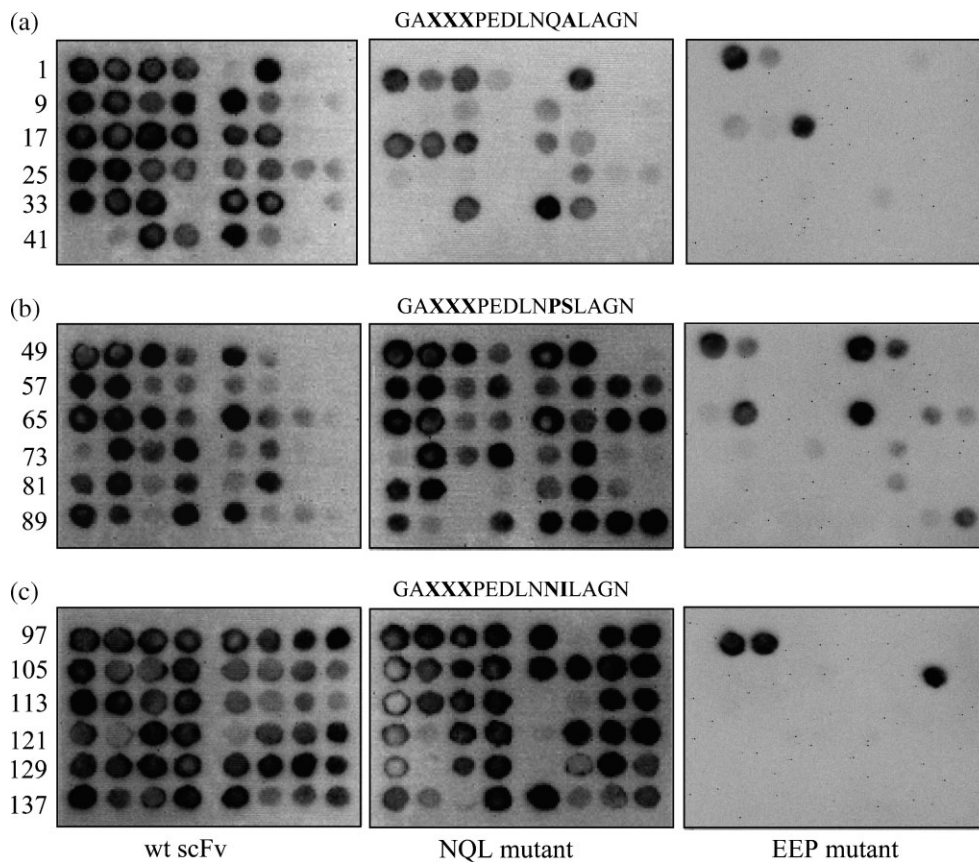


Figure 4 Binding of wt and mutated scFv to cellulose bound peptides substituted at positions 3–5 (*Lib 5*). The binding pattern of the substituted peptides based on the Lys¹⁰Ala modified h-peptide sequence in panel (a), based on the Gln⁹Pro/Lys¹⁰Ser modified h-peptide sequence in panel (b) and based on the Gln⁹Asn/Lys¹⁰Ile modified sequence in panel (c). Each membrane contained 48 peptides, where each peptide was synthesized as an individual spot on cellulose membrane.

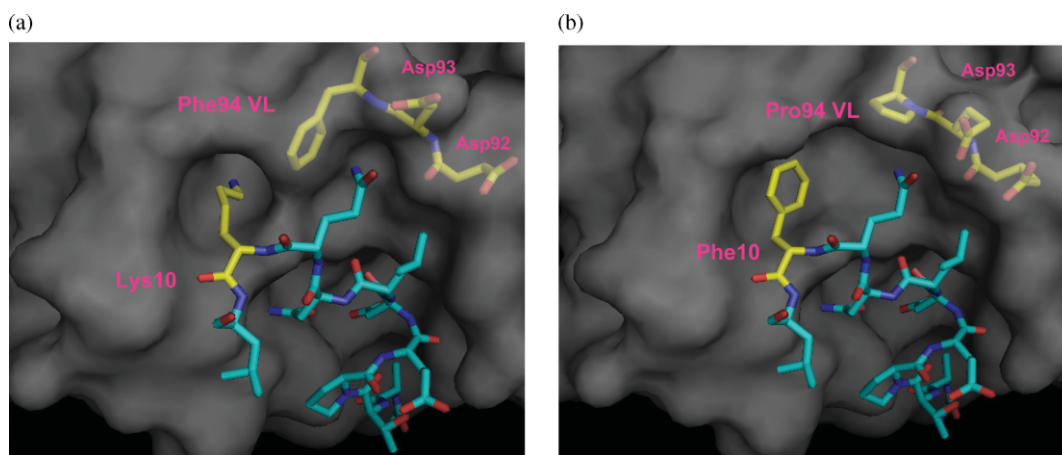


Figure 5 (a) Model for the interactions of the wt h-peptide with the wt CB4-1 Fab in the antibody binding site; coordinates taken from PDB 1BOG. (b) Model for the interactions of the Lys¹⁰Phe modified h-peptide with the EEP mutant antibody. The Lys¹⁰Phe peptide–EEP mutant antibody complex was built by homology modeling using the CB4-1 Fab–h-peptide complex (PDB 1BOG) as a template. Molecular surfaces of antibodies are shown in gray and the carbons of antigen residues in cyan. The carbons of the CB4-1 Fab and mutated antigen residues are colored in yellow. The other atoms of the residues are color coded in blue (nitrogen) and red (oxygen), Lys¹⁰ fits into the hydrophobic cavity formed by surrounding residues and makes a hydrophobic interaction with VL: Phe94 of the antibody (a). The hydrophobic cavity is partially distracted by VL: Phe94 replacement to Pro (b). Phe can easily be accepted in the peptide position without sterical hindrance. Pictures were made using PyMOL (DeLano Scientific).

longer and larger aromatic side chains, such as Phe, is presumably less favored for interaction with the wt antibody due to some steric restrictions. However, such substitutions improved the affinity of the EEP mutant compared to h-peptide (Table 6). Molecular modeling shows that this Phe substitution can easily be accepted without sterical hindrance when the antibody Phe⁹⁴ is replaced by the smaller Pro (Figure 5(b)).

A likely explanation for the increased selectivity of the h-peptide double mutant Gln⁹Pro/Lys¹⁰Ser toward the NQL mutant scFv is, probably, better hydrophobic interactions between Leu⁹⁴ of the antibody and the peptide's Pro⁹ or conformational changes of the peptide introduced by Pro, which resulted in a better contact with the NQL mutant scFv.

DISCUSSION

The identification of peptides that bind to antibodies is an important step in characterizing antibody specificity in studies of their molecular recognition. Developing antibodies as diagnostics, research tools, or therapeutics often require detailed information about their interaction with peptide antigens. A common approach is the generation and screening of chemical combinatorial libraries or phage/ribosome displays. Generally, such approaches rely on the generation of a huge number of random peptides that are screened for binders using more or less simple screening procedures. An alternative approach would be to select potential binders in a rational way.

In this study, we examined the potential of using SMD to investigate the requirements for antibody–antigen interactions. A common approach in SMD is to first describe the amino acids of the peptide selected for modification by molecular descriptors. Amino acid z -scales were used in our case and are a favorable choice as they are orthogonal and based on physicochemical properties and accordingly, therefore, can be used to identify chemically diverse sets of amino acids for substitutions. Moreover, we here used SMD to identify subsets of 6 and 10 amino acids from the 20 natural amino acids. These sets provide different levels of coverage of the chemical space, and when several small libraries are created from them in sequence they will result in smaller library sizes, potentially without loss of information gained. The numbers of peptides selected from these sets were then further reduced by D-optimal design that chooses a given number of peptides covering the largest structural volume of chemical space possible with a given number of peptides selected for a set [21].

Evaluations of the three initial libraries, which were based on modifications of the *N*-terminal core and *C*-terminal parts of the h-peptide, revealed the most critical residues for the binding of the antigen to the wt and mutated antibodies. Only a few binders were

obtained from the peptide libraries with modifications at positions 6–8 or 9–11; all of these contained the original leucines 7 and 11 of the h-peptide. These results essentially confirm earlier findings from substitutional analysis of the h-peptide's interaction with the original mCB4-1 antibody underlying the wt scFv of the present study. Thus, the earlier substitutional analysis revealed that the h-peptide strongly selects leucines at positions 7 and 11, which are obviously key binding residues, while aspartic acid 6 and asparagine 8 are less important for the CB4-1 antibody binding [17,29]. Besides, the mutations introduced in CDR L3 loop did not alter the obligatory requirement of leucine at positions 7 and 11 of the antigen.

The present study thus revealed that all the substituted peptides that showed some binding activity to both wt and mutated antibodies had the wt leucines in their sequences. The importance of these residues for interactions may be deduced from the crystal structure of the CB4-1 Fab/h-peptide complex. Thus, both the leucines are involved in direct contacts with the heavy and light chains residues in the CDR loops of the antibody, establishing essential hydrophobic contacts [23]. The results of the present study also indicated that substitutions at positions 6 and 8 of the antigen were more critical for affording binding to the NQL and EEP mutated antibodies compared to binding to the wt antibody. Such increased selectivity for modifications at these positions has been earlier observed for scFv CB4-1 single point mutant Phe⁹⁴Ala in the interaction with epitope derived peptides [30].

The design of the libraries that followed the initial ones was biased to keep key residues intact and modify the Gln⁹Lys¹⁰ residues that were less critical to substitutions. More binders were obtained for both the NQL and EEP mutants from the libraries created with the above design constraints, and these showed binding K_d s in the 10^{-6} – 10^{-8} M range. Several of these new peptides showed even slightly improved binding ability to the mutated scFvs compared to their binding to the wt antibody. Moreover, the h-peptide double substitution Gln⁹Pro/Lys¹⁰Ser increased the affinity of the antigen for the NQL mutant antibody, while it reduced it for the wt antibody. This might be due to improved hydrophobic interactions between Pro⁹ of the peptide and antibodies Leu⁹⁴ or some conformational changes introduced by Pro. Additional accumulation of further substitutions at the *N*-terminal end of h-peptide, already modified at positions 9 and 10, showed that this peptide can accommodate several substitutions without loss of binding activity.

Thus, applying SMD sequentially on libraries allowed us to identify important residues responsible for peptide antigen–antibody interactions and improved binding toward mutated antibodies with use of only small sized peptide libraries. On top of that, we were able

to find several peptides that were capable of binding slightly better to one of the mutant antibodies than to the wt antibody. It thus appears that SMD can be employed as an effective strategy to evaluate the importance of amino acids for interaction, and improve the affinity of low or moderate binders without the need for the manufacture of extensively large libraries. The advantages of SMD is thus that a small set of substitutions can be determined and directed libraries can be created much faster than when using extensively sized library methods, such as phage and ribosome display. Moreover, the rules learned from the studies applying SMD leads to better understanding of the relationship between sequence/structure and function, which eventually will allow even more rational design strategies to be developed in the future.

Acknowledgements

Supported by the Swedish VR-M (04X-05957).

REFERENCES

- Chothia C, Lesk AM. Canonical structures for the hypervariable regions of immunoglobulins. *J. Mol. Biol.* 1987; **196**: 901–917, DOI: 10.1016/0022-2836(87)90412-8.
- Braden BC, Poljak RJ. Structural features of the reactions between antibodies and protein antigens. *FASEB J.* 1995; **9**: 9–16.
- Appel JR, Buencamino J, Houghten RA, Pinilla C. Exploring antibody polyspecificity using synthetic combinatorial libraries. *Mol. Divers.* 1996; **2**: 29–34.
- Dionne SO, Lake DF, Grimes WJ, Smith MH. Identification of HLA-Cw6.02 and HLA-Cw7.01 allele-specific binding motifs by screening synthetic peptide libraries. *Immunogenetics* 2004; **56**: 391–398, DOI: 10.1007/s00251-004-0710-1.
- Hoogenboom HR, de Bruine AP, Hufton SE, Hoet RM, Arends JW, Roovers RC. Antibody phage display technology and its applications. *Immunotechnology* 1998; **4**: 1–20.
- Schaffitzel C, Hanes J, Jermutus L, Pluckthun A. Ribosome display: an in vitro method for selection and evolution of antibodies from libraries. *J. Immunol. Methods* 1999; **231**: 119–135, DOI: 10.1016/S0022-1759(99)00149-0.
- Slootstra JW, Puijk WC, Ligtoet GJ, Kuperus D, Schaaper WM, Meloen RH. Screening of a small set of random peptides: a new strategy to identify synthetic peptides that mimics epitopes. *J. Mol. Recognit.* 1997; **10**: 217–224.
- Reineke U, Ivascu C, Schlieff M, Landgraf C, Gericke S, Zahn G, Herzel H, Volkmer-Engert R, Schneider-Mergener J. Identification of distinct antibody epitopes and mimotops from a peptide array of 5520 randomly generated sequences. *J. Immunol. Methods* 2002; **267**: 37–51, DOI: 10.1016/S0022-1759(02)00139-4.
- Slootstra JW, Puijk WC, Ligtoet GJ, Langeveld JP, Meloen RH. Structural aspects of antibody–antigen interaction revealed through small random peptide libraries. *Mol. Divers.* 1996; **1**: 87–96.
- Linusson A, Wold S, Norden B. Statistical molecular design of peptoid libraries. *Mol. Divers.* 1999; **4**: 103–114, DOI: 10.1023/A:1026416430656.
- Andersson K, Choulier L, Hamalainen MD, van Regenmortel MH, Altschuh D, Malmqvist M. Predicting the kinetics of peptide–antibody interactions using a multivariate experimental design of sequence and chemical space. *J. Mol. Recognit.* 2001; **14**: 62–71, DOI: 10.1002/1099-1352(200101/02)14:1<62::AID-JMR520>3.0.CO;2-T.
- Choulier L, Andersson K, Hamalainen MD, van Regenmortel MH, Malmqvist M, Altschuh D. QSAR studies applied to the prediction of antigen–antibody interaction kinetics as measured by BIACORE. *Protein Eng.* 2002; **15**: 373–382.
- Dejaegere A, Choulier L, Lafont V, De Genst E, Altschuh D. Variations in antigen–antibody association kinetics as a function of pH and salt concentration: a QSAR and molecular modeling study. *Biochemistry* 2005; **44**: 14409–14418, DOI: 10.1021/bi050986v S0006-2960(05)00986-4.
- De Genst E, Areskoug D, Decanniere K, Muyldermans S, Andersson K. Kinetic and affinity predictions of a protein–protein interaction using multivariate experimental design. *J. Biol. Chem.* 2002; **277**: 29897–29907, DOI: 10.1074/jbc.M202359200.
- Van Regenmortel MH, Choulier L. Recognition of peptides by antibodies and investigations of affinity using biosensor. *Comb. Chem. High Throughput Screen.* 2001; **5**: 385–395.
- Mandrika I, Prusis P, Yahorava S, Tars K, Wikberg JE. QSAR of multiple mutated antibodies. *J. Mol. Recognit.* 2007; **20**: 97–102, DOI: 10.1002/jmr.817.
- Kramer A, Keitel T, Winkler K, Stocklein W, Hohne W, Schneider-Mergener J. Molecular basis for the binding promiscuity of an anti-p24 (HIV-1) monoclonal antibody. *Cell* 1997; **91**: 799–809, DOI: 10.1016/S0092-8674(00)80468-7.
- Linusson A, Gottfries J, Lindgren F, Wold S. Statistical molecular design of building blocks for combinatorial chemistry. *J. Med. Chem.* 2000; **43**: 1320–1328, DOI: 10.1021/jm991118x S0022-2623(99)01118-8.
- Lundstedt T, Seifert E, Abramo L, Thelin B, Nystrom A, Petersen J, Bergman R. Experimental design and optimization. *Chemometr. Intell. Lab. Syst.* 1998; **42**: 3–40.
- Sandberg M, Eriksson L, Jonsson J, Sjoström M, Wold S. New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. *J. Med. Chem.* 1998; **41**: 2481–2491, DOI: 10.1021/jm9700575.
- DeAguiar PF, Bourguignon B, Khots MS, Massart DL, Phan-Thau-Luu R. D-optimal designs. *Chemom. Intell. Lab. Syst.* 1995; **30**: 199–210.
- Leatherbarrow RJ. *GraFit Version 5*. Erithacus Software Ltd: Horley, 2001.
- Keitel T, Kramer A, Wessner H, Scholz C, Schneider-Mergener J, Hohne W. Crystallographic analysis of anti-p24 (HIV-1) monoclonal antibody cross-reactivity and polyspecificity. *Cell* 1997; **91**: 811–820, DOI: 10.1016/S0092-8674(00)80469-9.
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models and electron density maps and the location of errors in these models. *Acta Crystallogr.* 1991; **47**: 110–119.
- Lindahl E, Azuara C, Koehl P, Delarue M. NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. *Nucleic Acids Res.* 2006; **34**: W52–W56, DOI: 10.1093/nar/gkl082.
- DeLano WL. *The PyMOL Molecular Graphics System*. DeLano Scientific: San Carlos, CA, 2002.
- Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C. *Sequences of Proteins of Immunological Interest*, 5th edn. Public Health Service. National Institutes of Health: Cambridge, 1991.
- Frank R. Spot synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. *Tetrahedron* 1992; **48**: 9217–9232.
- Hoffmüller U, Knaute T, Hahn M, Höhne W, Schneider-Mergener J, Kramer A. Evolutionary transition pathways for changing peptide ligand specificity and structure. *EMBO J.* 2000; **19**: 4866–4874, DOI: 10.1093/emboj/19.18.4866.
- Winkler K, Kramer A, Kuttner G, Seifert M, Scholz C, Wessner H, Schneider-Mergener J, Hohne W. Changing the antigen binding specificity by single point mutations of an anti-p24 (HIV-1) antibody. *J. Immunol.* 2000; **165**: 4505–4514.