Evolutionary transition pathways for changing peptide ligand specificity and structure

Ulrich Hoffmüller¹, Tobias Knaute¹,
Michael Hahn², Wolfgang Höhne²,
Jens Schneider-Mergener^{1,3} and
Achim Kramer^{1,4}

¹Institut für Medizinische Immunologie and ²Institut für Biochemie, Universitätsklinikum Charité, Humboldt-Universität zu Berlin, Schumannstrasse 20/21, 10098 Berlin, Germany

⁴Present address: Department of Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, MA 02115, USA

³Corresponding author
e-mail: jsm@charite.de

We identified evolutionary pathways for the inter-conversion of three sequentially and structurally unrelated peptides, GATPEDLNQKL, GLYEWGGARI and FDKEWNLIEQN, binding to the same site of the hypervariable region of the anti-p24 (HIV-1) monoclonal antibody CB4-1. Conversion of these peptides into each other could be achieved in nine or 10 single amino acid substitution steps without loss of antibody binding. Such pathways were identified by analyzing all 7 620 480 pathways connecting 2560 different peptides, and testing them for CB4-1 binding. The binding modes of intermediate peptides of selected optimal pathways were characterized using complete sets of substitution analogs, revealing that a number of sequential substitutions accumulated without changing the pattern of key interacting residues. At a distinct step, however, one single amino acid exchange induces a sudden change in the binding mode, indicating a flip in specificity and conformation. Our data represent a model of how different specificities, structures and functions might evolve in protein–protein recognition.

Keywords: molecular evolution/peptide library/spot synthesis/transition pathways

Introduction

The specificity of molecular recognition events selected during evolution is an essential component in the development of highly complex and differentiated physiological processes in living organisms. Protein–protein interactions play an essential role in the immunological discrimination of pathogens, active transport of proteins between organelles and signal transduction events, in addition to being important in regulatory and structural networks. However, this stringent specificity is not absolute, as has become increasingly obvious in the past few years through analyzing the specificity of B and T cell immune reactions. Indeed, screening of biologically or chemically generated peptide libraries have shown that monoclonal antibodies and also T cells are able to recognize peptide antigens that

are completely unrelated in sequence and structure (Kramer *et al.*, 1997 and literature cited therein; Pinilla *et al.*, 1998; Grogan *et al.*, 1999; DeLano *et al.*, 2000; Maier *et al.*, 2000). Binding promiscuity has also been observed for proteins other than antibodies (Sugamura *et al.*, 1996; Wrighton *et al.*, 1996).

The monoclonal anti-p24 (HIV-1) antibody CB4-1 (Grunow *et al.*, 1990; Höhne *et al.*, 1993) used here is able to recognize specifically several peptide antigens unrelated in sequence and structure (Kramer *et al.*, 1997). The original 11meric p24 epitope GATPQDLNML or a higher affinity homolog GATPEDLNQKL (h-pep) recognizes CB4-1 in an extended conformation. The unrelated peptide GLYEWGGARI (u1-pep), however, has a loop structure in its CB4-1 binding conformation (Figure 1; Keitel *et al.*, 1997). Furthermore, although the two peptides bind to the same cleft formed by the light and heavy hypervariable region of the antibody, the key interacting residues in the peptides and antibody are different (Keitel *et al.*, 1997; Kramer *et al.*, 1997). For a second unrelated CB4-1 binding peptide FDKEWNLIEQN (u2-pep) no crystal structure for the peptide–antibody complex is yet available, but a model is being proposed by soft docking techniques (Figure 1). These peptides were identified by testing chemically generated libraries of $>10^5$ different peptide mixtures or peptides (Kramer *et al.*, 1997), which were prepared in a positionally defined manner on continuous cellulose membranes by spot synthesis (Frank, 1992; Kramer and Schneider-Mergener, 1998).

The observation that a protein, in this case an antibody, is able to recognize unrelated sequences and structures led us to the question of whether these specificities and structures can be converted into each other. Such an investigation would not only provide insights into the molecular evolution of protein–protein recognition events but could also address one of the key questions in evolution: whether changes in phenotype are continuous or saltatory. While the gradual changes in the genotype imply gradual changes in phenotype as well, saltatory changes explain better how evolutionary gaps can be bridged. With respect to the evolution of protein–ligand interactions, peptides are the smallest units possessing a phenotype represented at three different levels: sequence, structure and function.

To answer the question of whether the different CB4-1 peptide ligands can be reciprocally converted into each other, we analyzed all possible single step transition pathways ($>7.6 \times 10^6$) between the three ligands. Using the software PepTrans (Hoffmüller and Schneider-Mergener, 1999), which we designed for this purpose, we synthesized all 2560 possible peptides and performed binding studies (for calculations see Results). To understand better the structural mechanisms involved in the

Fig. 1. The antibody CB4-1 complex with binding peptides. Structures of the peptides GATPEDLNQKL (h-pep, blue), GLYEWGGARI (u1-pep, red) and FDKEWNLIEQN (u2-pep, green) in a complex with the monoclonal antibody CB4-1 (gray) are superimposed. The structures of h-pep and u1-pep were obtained by X-ray crystallography (Keitel *et al.*, 1997). Internal Coordinate Mechanics (ICM, Abagyan *et al.*, 1994) was used to generate a model of the u2-pep conformation as well as for visualization of the complexes.

transformation of sequence and structure we synthesized complete sets of substitution analogs by exchanging each residue (by all other 19 amino acids) of all the peptide intermediates from each best transition pathway, and analyzed them for CB4-1 binding. Furthermore, one transition pathway was investigated for binding to two CB4-1 single chain mutants to study the effects of accompanying changes in the other binding partner.

Results

Testing all transition pathways between CB4-1 binding peptides

To find out whether inter-conversion pathways exist for the stepwise conversion of peptides GATPEDLNQKL (h-pep), GLYEWGGARI (u1-pep) and FDKEWNLIEQN (u2-pep) into each other by single amino acid substitutions we analyzed all possible transition intermediates for antibody CB4-1 binding. The transformation of GATPEDLNQKL into DGLYEWGGARI (h-pep $\leftrightarrow$ u1-pep, the N-terminal aspartic acid added for length adjustment does not influence binding) or into FDKEWNLIEQN (h-pep $\leftrightarrow$ u2-pep) involves $2^{10} = 1024$ possible peptides (two 11meric peptides, 10 different residues, positions 4 and 7, respectively, are identical; Table I). To convert GLYEWGGARIT (T, which has no effect on binding was added for length adjustment) into FDKEWNLIEQN (u1-pep $\leftrightarrow$ u2-pep) there are $2^9 = 512$ possible intermediate peptides (two 11meric peptides, nine different residues, positions 4 and 5 are identical). The sequences for these 2560 peptides (total number) were generated using the software PepTrans (Hoffmüller and Schneider-Mergener, 1999). The peptides were synthesized by spot synthesis, tested for CB4-1 binding (Figure 2), and the binding intensities were quantified (Table I). Roughly 7–8% of the peptide intermediates showed clear CB4-1 binding.

Identification of optimal transition pathways

The PepTrans software was then used to analyze the data from the binding studies (Figure 2 and Table I) to investigate whether successive point ‘mutation’ pathways, without loss of binding at any step, exist between the unrelated peptides. The procedure is illustrated using the transformation h-pep $\leftrightarrow$ u1-pep as an example (Table I). PepTrans tested all 10 factorial ($10!$, 3 628 800) pathways for this transformation (GATPEDLNQKL into DGLYEWGGARI; 11meric peptides, 10 different residues) and judged their quality with reference to their weakest binding peptide intermediate. Since several different pathways involving no loss of binding were identified, the one with the highest total binding intensity for all 11 intermediate peptides was chosen for further analysis. All 11 transition peptides of this pathway were synthesized as soluble peptides and dissociation constants of the respective peptide–antibody complexes were determined (Figure 3A). The same analyses were performed for the conversions u1-pep $\leftrightarrow$ u2-pep and h-pep $\leftrightarrow$ u2-pep (Figure 3B and C).

The results demonstrate that each peptide can be transformed into a completely unrelated peptide without loss of biologically relevant binding affinity on the way (Cumano and Rajewsky, 1986; Sharon, 1990). This study also revealed that several of the mutations that accumulated did not significantly affect affinity.

Binding of antibody CB4-1 mutants to h-pep $\leftrightarrow$ u1-pep intermediate peptides

Evolution of protein–protein recognition processes leading into new specificities and functions necessitates variation affecting both interacting partners. Therefore, we also tested the binding behavior of peptide intermediates in the h-pep $\leftrightarrow$ u1-pep transition pathway with respect to two very similar proteins, i.e. single chain antibody CB4-1 mutants V_L:Phe94Ala and V_H:Tyr32Ala (Kramer *et al.*,

Table I. Intermediate peptide sequences and the corresponding binding signals

Spot no.	Sequence	intensity
1	GATPEDLNQKL	337246
2	DATPEDLNQKL	336571
3	GGTPEDLNQKL	301347
4	DGTPEDLNQKL	311988
5	GALPEDLNQKL	307521
6	DALPEDLNQKL	293588
7	GGLPEDLNQKL	266846
8	DGLPEDLNQKL	268345
9	GATYEDLNQKL	158756
10	DATYEDLNQKL	126227
11	GGTYEDLNQKL	104054
12	DGTYEDLNQKL	99862
13	GALYEDLNQKL	140583
14	DALYEDLNQKL	102586
15	GGLYEDLNQKL	119743
16	DGLYEDLNQKL	119089
17	GATPEWLNQKL	217300
18	DATPEWLNQKL	228076
19	GGTPEWLNQKL	173608
20	DGTPEWLNQKL	161231
21	GALPEWLNQKL	169187
22	DALPEWLNQKL	201812
23	GGLPEWLNQKL	133710
24	DGLPEWLNQKL	133606
25	GATYEWLNQKL	77191
26	DATYEWLNQKL	95020
27	GGTYEWLNQKL	94667
28	DGTYEWLNQKL	83878
29	GALYEWLNQKL	55610
30	DALYEWLNQKL	76542
31	GGLYEWLNQKL	50777
...	...	...
...	...	...
1024	DGLYEWGGARI	61965

The table shows the sequences of the intermediate peptides of the h-pep (GATPEDLNQKL, blue) into u1-pep (DGLYEWGGARI, red) transformation and the corresponding spot numbers as well as the signal intensities of the CB4-1 binding studies. Amino acids are colored according to the peptide they derive from. Glutamic acid (E, black) is identical in both peptides.

1997; Figure 3D). The u1-pep and adjacent peptides were almost exclusively recognized by the mutant V_H:Tyr32Ala whereas there is a clear preference of the V_L:Phe94Ala mutant for h-pep and its nearest relatives. One substitution in u1-pep even led to higher binding affinity of the CB4-1 V_H:Tyr32Ala mutant. Here again, either peptide can accumulate some substitutions without significant loss of binding.

Key interacting residues of peptide intermediates reveal binding mode switch

The inter-conversion of the two sequentially unrelated peptides h-pep and u1-pep involves a complete change in structure of the bound ligand (Figure 1; Keitel *et al.*, 1997).

Fig. 2. Binding of CB4-1 to transition peptide analogs. All possible intermediate peptide sequences for the direct stepwise conversion of h-pep ↔ u1-pep (2¹⁰, 11mer peptides, 10 different residues), h-pep ↔ u2-pep (2¹⁰) and u1-pep ↔ u2-pep (2⁹) were synthesized on cellulose membranes and tested for CB4-1 binding (three different formats). All intermediates were generated using the PepTrans software as shown in Table I.

One way to obtain vital information about the binding mode and thus the structure of a peptide binding to an antibody is through substitutional analysis. In this analysis each residue of a peptide ligand is substituted by all other 19 L-amino acids, thus identifying key interacting residues (Kramer *et al.*, 1997; Reineke *et al.*, 1999). To investigate the change in binding modes during peptide transmutation we synthesized complete sets of substitution analogs for the h-pep ↔ u1-pep and the two other transition pathways.

As an example, in h-pep, leucines 7 and 11 are key residues for binding, while alanine 2, aspartic acid 6 and asparagine 8 are less important (Figure 4, left column). In u1-pep, leucine 3, tryptophan 6, glycines 7 and 8 and isoleucine 11 are key residues. Clearly, the transition started without affecting these key residues, the newly introduced amino acids proving to be unimportant. However, at a distinct step the key residue pattern completely changes into the essential pattern of the other peptide. For instance, peptide DATPEWLGARL (step 5) has five h-pep- and five u1-pep-derived residues (ignoring glutamic acid). Its key residue pattern, however, is identical to that of h-pep. The two leucine residues remained essential for binding (Figure 4, compare intensities with wt column). Stronger spots due to substitutions in either direction of transition indicate higher affinity binding analogs, such as glycine 8 to asparagine or tryptophan 6 to aspartic acid. The adjacent peptide DATPEWGGARL (step 6) with six u1-pep-derived residues has a pattern resembling u1-pep, with WGG now proving to be key residues. Substitution in either direction again leads to a significant increase in affinity, for example leucine 11 to isoleucine and glycine 7 to leucine (see also Figure 3).

The same phenomena were observed for the other transition pathways as qualitative comparison of the spot patterns in the other two columns reveal (Figure 4). Again, several substitutions can be accumulated before a sudden change in the binding pattern of the substitution analogs

recognition (Grogan *et al.*, 1999; Maier *et al.*, 2000). Furthermore, the involvement of polyspecificity and cross reactivity in biological recognition events has been implicated in the onset of autoimmune diseases (Oldstone, 1998), although this is still little understood at the molecular level. Here, we provide an example of how peptide ligands can be converted into sequentially and structurally unrelated molecules binding to the same ligate, in this case a monoclonal antibody. All possible direct transition pathways could be analyzed because in our model system the different ligands were known and are relatively small molecules. This enabled us to investigate a clearly defined molecular example of how different specificities and structures can evolve in protein–protein recognition. Several single step transition pathways between three unrelated molecules involving no loss of binding function were revealed (Figure 3; other pathways were analyzed quantitatively but not shown) and thus it seems likely that the ability to transform ligands is a general principle in protein–protein recognition. Direct transformation paths including simultaneous double substitutions consist of the same intermediate sequences and are therefore included in the experiments described. Since multiple simultaneous amino acid exchanges in close proximity are very unlikely, such processes do not play an important role in evolution and were not regarded in this study.

Molecular evolution of new functions and specificities requires changes in both binding partners and this was demonstrated by binding studies with two antibody point mutants. In the example shown, starting from the mid-point intermediate peptide DATPEWLGARL (Figure 3A), which is recognized only by wild-type CB4-1, the

equivalent of simultaneous substitutions in both binding partners (i.e. transition peptides in both directions and antibody mutants) led to two independent types of specific interaction with one or other of the mutants (Figures 3D and 6). These new binding partner pairs could represent starting points for transitions to as yet unknown specificities. Moreover, binding promiscuity between proteins might even facilitate evolutionary events in molecular recognition processes.

It is practically very difficult to elucidate the structures of all intermediates in a complex with CB4-1 by crystallography or NMR due to the low affinities observed at the critical stages of transition. However, substitutional analyses provide a clear, although lower resolution picture, representing a fingerprint of a peptide's binding conformation (Reineke *et al.*, 1999). This was affirmed by the crystal structure of h-pep $\leftrightarrow$ u1-pep intermediate 4 in complex with CB4-1. Such analyses of the intermediate peptides to monitor the changes in binding and how the transition takes place structurally revealed a common scheme for all three pathways analyzed. The binding conformations remained constant for several successive substitutions and then suddenly flipped into completely distinct conformations that underwent no further change in the remaining steps.

In general, the molecular phenotype of the peptides can be viewed from three different levels: (i) amino acid sequence, (ii) conformation of the peptide in a complex with the antibody and (iii) drift of potential function, here, altered binding profiles to CB4-1 point mutants. The h-pep $\leftrightarrow$ u1-pep transition data can therefore be summarized schematically as shown in Figure 6. The two distinct spots in conformational space, above, correspond

Fig. 5. Complex structure of h-pep and the transition analog DATPEDLGARL. X-ray structures of the peptides GATPEDLNQKL (h-pep, blue) and DATPEDLGARL (h-pep $\leftrightarrow$ u1-pep intermediate 4, brown) with CB4-1 hypervariable regions (gray). For visualization the program ICM (Figure 1) was used.

Fig. 4. Substitutional analyses of peptide intermediates of the three best transition pathways. Complete sets of single substitution analogs of each intermediate were prepared on cellulose membranes and tested for CB4-1 binding. Each position of the peptides was substituted by 20 amino acids (rows). The spots in the first column are identical and correspond to the starting peptides (wt, control). Other spots are single substitution analogs. Binding intensities can only be compared quantitatively within one substitutional analysis due to different incubation reactions and exposure times.

Fig. 6. Diagram correlating sequence, structure and function. The central level represents the sequence space in which h-pep (blue), u1-pep (red) and the respective transition analogs (different shades of purple) are located. The transition intermediates differ in only one amino acid and hence are located next to each other, the unshaded neighboring circles depicting alternative changes and possible routes between red and blue. The optimal pathway is where the amino acid sequence is altered continuously by small steps while retaining the function of CB4-1 binding. When bound to CB4-1 each peptide takes up a distinct position in conformational space (top level) as indicated. Two different and therefore distant clusters of conformations exist (compare Figures 1 and 4). The conformations in either cluster are very similar to each other (close in conformational space) and to one or other of the starting peptides. In the functional space (bottom level) three different positions are defined by peptide binding to CB4-1 (purple) and the two CB4-1 mutants (blue and red). Each peptide possesses one (wt binding) or more function(s) indicated by lines linking the sequence space and the functional space.

to the different starting peptide-related binding modes adopted by the intermediate peptides, which form a continuous path in the sequence space. The differential binding behavior of CB4-1 and either of its two mutants is illustrated by the sequence space peptides converging on one or two of the three spots in the functional space.

A key question in studying evolution is whether changes in phenotype are gradual or saltatory. Our results provide molecular evidence that mutations can accumulate in preparation for a sudden change in structure and function.

Analogous observations were made correlating changes of sequence and structure in RNA molecules (Fontana and Schuster, 1998). The evolutionary usefulness of accumulating neutral mutations preparing for saltatory changes is underlined by the recent discovery that a heat shock protein (hsp90) suppresses the phenotypic appearance of morphological mutations in *Drosophila* under normal conditions. By masking a hidden reservoir of genetic diversity this facilitates saltatory morphological changes upon environmental change (Rutherford and Lindquist, 1998).

Materials and methods

Peptide synthesis

Cellulose-bound peptides and peptide mixtures were prepared by semi-automated spot synthesis (Frank, 1992; Abimed, Langenfeld, Germany; Software LISA, Jerini BioTools GmbH, Berlin, Germany) using Whatman 50 (Whatman, Maidstone, UK) cellulose membranes as described before in detail (Kramer *et al.*, 1994; Kramer and Schneider-Mergener, 1998). All peptides were acetylated N-terminally.

The peptides for the solution phase binding experiments (ELISA and BIAcore measurements) with CB4-1 were prepared according to standard Fmoc machine protocols using a multiple peptide synthesizer (Abimed, Langenfeld, Germany) and analyzed by reverse-phase HPLC and MALDI-TOF mass spectrometry. N-terminal acetylation was carried out using acetic anhydride and diisopropylethylamine; biotinylation using biotin (Fluka, Neu-Ulm, Germany) activated with benzotriazole-1-yl-oxy-tripyrrolidinophosphonium hexafluorophosphate (PyBOP, Novabiochem) and *N*-methylmorpholine. The N-terminally biotinylated peptides contained two β -alanine residues as a spacer between biotin and the peptide.

Antibody binding studies with cellulose-bound peptides

The membrane-bound libraries were blocked overnight with blocking buffer, i.e. blocking reagent (CRB, Norwich, UK) in T-TBS (Tris-buffered saline–0.05% Tween 20) containing 1% sucrose. After washing with T-TBS, 0.1 μ g/ml CB4-1 in blocking buffer was added and incubated for 3 h at room temperature. After washing three times with T-TBS, a peroxidase-labeled anti-mouse antibody (Sigma, Munich, Germany; 1 μ g/ml in blocking buffer) was applied for 2 h at room temperature. Detection used a chemiluminescence substrate and imager (Boehringer Mannheim, Mannheim, Germany).

Competitive ELISA studies

N-terminally acetylated peptides were tested for their ability to compete with the CB4-1–p24 interaction in solution. Microtitre plates (Nunc, Roskilde, Denmark) were coated with 0.1 μ g/ml recombinant p24 (Hausdorf *et al.*, 1994) in 0.1 M sodium carbonate buffer pH 9.6 and incubated for 20 h at 4°C. After washing three times with phosphate-buffered saline (PBS)–0.1% Tween 20, 0.1 μ g/ml horseradish peroxidase-labeled CB4-1 was added with peptides in various concentrations (depending on the respective inhibition constants) in PBS–0.1% Tween 20 containing 6% Gelifundol®S (Biotest, Dreieich, Germany) in a total volume of 50 μ l for 20 h at 4°C. After washing three times with PBS–0.1% Tween 20, the bound enzymic activity was determined by adding 5.5 mM *o*-phenylenediamine hydrochloride (Fluka, Buchs, Switzerland) and 8.5 mM H₂O₂ in 0.1 M citrate buffer pH 5.0. The reaction was terminated after 10 min by adding 1 M sulfuric acid containing 0.05 M sodium sulfite. The absorbance was measured at 492 and 620 nm, as reference, using an ELISA reader (Anthos, Köln, Germany). Inhibition constants were calculated according to Friguet *et al.* (1985).

Affinity measurements with BIAcore

Equipment and reagents. BIAcore 1000 system, sensor chips CM5, buffer HBS (10 mM HEPES with 0.15 M NaCl, 3.4 mM EDTA and 0.005% surfactant P20 at pH 7.4), amine coupling kit and BIA evaluation software were obtained from BIAcore AB (Uppsala, Sweden).

Immobilization. Monoclonal antibody CB4-1 single chains were immobilized on CM5 chips using the amine coupling procedure described in the BIA application handbook (O'Shannessy and Wilchek, 1990). The amount of immobilized CB4-1 single chains corresponded to an increase in the SPR signal of ~5000 resonance units (RU) for flowcell 2. Appropriate amounts of non-specific monoclonal control single chain antibody 1F9 were immobilized in flowcell 1 as reference.

Binding experiments. All binding experiments were performed at 25°C with a flow rate of 25 μ l/min (injection volume 70 μ l). Peptides were used at various concentrations between 1 nM and 200 μ M. Complete regeneration was obtained after dissociation without using regeneration buffer.

Data analysis. Transformation of data and analysis were performed with BIA evaluation software. The control sensorgram (flowcell 1) was subtracted from the sensorgrams obtained with flowcell 2. The steady-state values of the binding equilibrium were plotted versus the different peptide concentrations and fitted using the implemented steady-state evaluation resulting in the dissociation constants for the antibody–peptide complexes.

PepTrans software

Algorithms for two tasks were written: (i) the generation of the sequences of the intermediate peptides as input for the pipetting robots to create the intermediate peptide libraries and (ii) the identification of possible transformation paths using the intermediate peptide sequences and their corresponding binding intensities.

For the generation of the intermediate peptide sequences, a list of all possible sequences was compiled, in which the amino acid at each position matches the corresponding residue of either the starting peptide or the end peptide sequence. Table I (second column) shows the beginning of the sequence list for the transformation h-pep $\leftrightarrow$ u1-pep. In this example, the amino acids of h-pep and u1-pep differ in 10 positions; therefore, the number of different intermediate sequences is $2^{10} = 1024$.

The search for transformation pathways including only binding peptides created by single amino acid substitutions was based on the measured CB4-1 binding intensities of the intermediate peptides (Figure 2). The algorithm is explained using the transformation h-pep $\leftrightarrow$ u1-pep as an example. The binding signals were quantified using a Lumi-Imager (Boehringer Mannheim) and linked to the sequence data, as shown in column 3 of Table I. Next, all possible direct transformation pathways containing only single amino acid substitutions were systematically checked, all varying in the order of substitution of the different positions. The first transformation pathway had the order of substitutions (1,2,3,4,5,6,7,8,9,10), meaning that the first substitution was the exchange of G1 by D, the second A2 by G, the third T3 by L and so on. The second transformation pathway had the order (2,1,3,4,5,6,7,8,9,10), meaning that first, A2 is substituted by G, then G1 by D, etc. In this manner, all permutations of the substitution orders were analyzed, a total of 10! (3.7 million) transformation pathways. For every single pathway, PepTrans generated the peptide sequence for each step, read the corresponding binding intensity in the sequence table and calculated the binding intensity minimum. Of the 3.7 million transformation pathways the program identified 50 transformation pathways with the 'highest binding intensity minima'. Of these identified pathways, the one with the highest sum of binding intensities of its intermediates was defined as the optimal pathway.

X-ray analysis of CB4-1–h-pep $\leftrightarrow$ u1-pep intermediate 4 complex

The complex of CB4-1 Fab fragment with transformation analog DATPEDLGARL (h-pep $\leftrightarrow$ u1-pep intermediate 4) was concentrated to 10 mg/ml and crystallized in a hanging drop set up at room temperature over a reservoir containing 1.8 M ammonium sulfate in 0.1 M MOPS buffer pH 7.5. Data were collected at beamline X11 at the EMBL Outstation at DESY, Hamburg, with nitrogen cooling. A data completeness of 99.5% was obtained up to a resolution of 2.6 Å. The structure was solved with AMoRe (Navaza, 1994) using the model from Keitel *et al.* (1997) and refined with refmac (Murshudov *et al.*, 1997). The current R_{free} is 33% for the complete model with 87 water molecules added.

Molecular modeling of the CB4-1–u2-pep complex

The program ICM (Internal Coordinate Mechanics; Abagyan *et al.*, 1994) was used to dock the u2-pep into the antibody according to a two-step procedure described elsewhere (Totrov and Abagyan, 1994; Zhou and Abagyan, 1998). In the first step, the totally flexibilized peptide was subjected to an exhaustive Monte Carlo conformational search with the antibody represented by a set of potential maps. Four independent calculations converged after ~63 million function calls into structurally and energetically similar conformations. In the second step, the structures were refined by sampling the side chains of both the peptide and the antibody binding pocket, which was now present as a full-atom model. In the refinement step the ECEPP/3 force field was supplemented by a surface term and a term representing the configurational entropy of side chains (Abagyan and Totrov, 1994).

Acknowledgements

We thank Karsten Winkler for providing antibody CB4-1 scFv mutants; Berit Hoffmann, Lying Dong, Christiane Landgraf and Ines Kretzschmar for technical assistance in peptide synthesis and binding studies. We would also like to thank the staff at EMBL Outstation at DESY, Hamburg for beamtime allocation and continuous support in the project (project number PX99-309). This work was supported by grants of the DFG, Fonds der Chemischen Industrie and Charité.

References

- Abagyan, R.A. and Totrov, M.M. (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.*, **235**, 983–1002.
- Abagyan, R., Totrov, M. and Kuznetsov, D. (1994) ICM—a new method for protein modelling and design: applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.*, **15**, 488–506.
- Cumano, A. and Rajewsky, K. (1986) Clonal recruitment and somatic mutation in the generation of immunological memory to the hapten NP. *EMBO J.*, **5**, 2459–2468.
- DeLano, W.L., Ultsch, M.H., de Vos, A.M. and Wells, J.A. (2000) Convergent solutions to binding at a protein–protein interface. *Science*, **287**, 1279–1283.
- Fontana, W. and Schuster, P. (1998) Continuity in evolution: on the nature of transitions. *Science*, **280**, 1451–1455.
- Frank, R. (1992) Spot synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. *Tetrahedron*, **48**, 9217–9232.
- Friguet, B., Chaffotte, A.F., Djavadi-Ohanian, L. and Goldberg, M.E. (1985) Measurements of the true affinity constant in solution of antigen–antibody complexes by enzyme-linked immunosorbent assay. *J. Immunol. Methods*, **77**, 305–319.
- Grogan, J.L., Kramer, A., Nogai, A., Dong, L., Ohde, M., Schneider-Mergener, J. and Kamradt, T. (1999) Crossreactivity of MBP-specific T cells with multiple microbial peptides: EAE-induction in TCR transgenic mice. *J. Immunol.*, **163**, 3764–3770.
- Grunow, R., Giese, R., Porstmann, T., Döpel, H., Hänsel, K. and von Baehr, R. (1990) Development and biological testing of human and murine monoclonal antibodies against HIV antigens. *Z. Klin. Med.*, **45**, 367–369.
- Hausdorf, G., Gewiss, A., Wray, V. and Porstmann, T.J. (1994) A recombinant human immunodeficiency virus type-1 capsid protein (rp24): its expression, purification and physico-chemical characterization. *Virol. Methods*, **50**, 1–9.
- Hoffmüller, U. and Schneider-Mergener, J. (1999) <http://www.charite.de/immunologie/peptrans>
- Höhne, W.E. *et al.* (1993) Structural base of the interaction of a monoclonal antibody against p24 of HIV-1 with its peptide epitope. *Mol. Immunol.*, **30**, 1213–1221.
- Keitel, T., Kramer, A., Wessner, H., Scholz, C., Schneider-Mergener, J. and Höhne, W. (1997) Crystallographic analysis of anti-p24 (HIV-1) monoclonal antibody cross-reactivity and polyspecificity. *Cell*, **91**, 811–820.
- Kramer, A. and Schneider-Mergener, J. (1998) Synthesis and screening of peptide libraries on cellulose membrane supports. *Methods Mol. Biol.*, **87**, 25–39.
- Kramer, A., Schuster, A., Reineke, U., Malin, R., Volkmer-Engert, R., Landgraf, C. and Schneider-Mergener, J. (1994) Combinatorial cellulose-bound peptide libraries: screening tool for the identification of peptides that bind ligands with predefined specificity. *Methods Comp. Methods Enzymol.*, **6**, 388–395.
- Kramer, A., Keitel, T., Winkler, K., Stöcklein, W., Höhne, W. and Schneider-Mergener, J. (1997) Molecular basis for the binding promiscuity of an anti-p24 (HIV-1) monoclonal antibody. *Cell*, **91**, 799–809.
- Kramer, A., Reineke, U., Dong, L., Hoffmann, B., Hoffmüller, U., Winkler, D., Volkmer-Engert, R. and Schneider-Mergener, J. (1999) Spot synthesis: observations and optimizations. *J. Peptide Res.*, **54**, 319–327.
- Maier, B., Molinger, M., Cope, A.P., Fugger, L., Schneider-Mergener, J., Sonderstrup, G., Kamradt, T. and Kramer, A. (2000) Multiple cross-reactive self-ligands for *Borrelia burgdorferi*-specific HLA-DR4-restricted T cells. *Eur. J. Immunol.*, **30**, 448–457.
- Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D*, **53**, 240–255.
- Navaza, J. (1994) AMoRe—an automated procedure for molecular replacement. *Acta Crystallogr. A*, **50**, 157–163.
- Oldstone, M.B. (1998) Molecular mimicry and immune-mediated diseases. *FASEB J.*, **12**, 1255–1265.
- O'Shannessy, D.J. and Wilchek, M. (1990) Immobilization of glycoconjugates by their oligosaccharides: use of hydrazo-derivatized matrices. *Anal. Biochem.*, **191**, 1–8.
- Pinilla, C., Appel, J.R., Campbell, G.D., Buencamino, J., Benkirane, N., Muller, S. and Greenspan, N.S. (1998) All-D peptides recognized by an anti-carbohydrate antibody identified from a positional scanning library. *J. Mol. Biol.*, **283**, 1013–1025.
- Reineke, U., Kramer, A. and Schneider-Mergener, J. (1999) Antigen sequence- and library-based mapping of linear and discontinuous protein–protein interaction sites by spot synthesis. *Curr. Top. Microbiol. Immunol.*, **243**, 23–36.
- Rutherford, S.L. and Lindquist, S. (1998) Hsp90 as a capacitor for morphological evolution. *Nature*, **396**, 336–342.
- Sharon, J. (1990) Structural correlates of high antibody affinity: three engineered amino acid substitutions can increase the affinity of an anti-*p*-azophenylarsenate antibody 200-fold. *Proc. Natl Acad. Sci. USA*, **87**, 4814–4817.
- Sugamura, K., Asao, H., Kondo, M., Tanaka, N., Ishii, N., Ohbo, K., Nakamura, M. and Takeshita, T. (1996) The interleukin-2 receptor γ chain: its role in the multiple cytokine receptor complexes and T cell development in XSCID. *Annu. Rev. Immunol.*, **14**, 179–205.
- Totrov, M.M. and Abagyan, R.A. (1994) Detailed *ab initio* prediction of lysozyme–antibody complex with 1.6 Å accuracy. *Nature Struct. Biol.*, **1**, 259–263.
- Wrighton, N.C. *et al.* (1996) Small peptides as potent mimetics of the protein hormone erythropoietin. *Science*, **273**, 458–464.
- Zhou, Y. and Abagyan, R.A. (1998) How and why phosphotyrosine-containing peptides bind to the SH2 and PTB domains. *Fold. Des.*, **3**, 513–522.

Received April 17, 2000; revised July 21, 2000;
accepted July 25, 2000