Conformations of Peptide Fragments Comprising the Complete Sequence of Component III of *Chi t* I and Their Relationship to T-Cell Stimulation†

Michael Czisch,‡,§ Verena Liebers,§ Reimond Bernstein,‡ Zhiping Chen,§ Xaver Baur,§ and Tad A. Holak*,‡

Max-Planck-Institut für Biochemie, D-82152 Martinsried bei München, FRG, and Berufsgenossenschaftliches Forschungsinstitut für Arbeitsmedizin an der Ruhr-Universität Bochum, Gilsingstrasse 14, D-44789 Bochum, FRG

Received March 14, 1994; Revised Manuscript Received June 8, 1994®

ABSTRACT: Conformational preferences of synthetic peptides that span the complete sequence of *Chironomus thummi* hemoglobin (*Chi t* I) component III were studied by nuclear magnetic resonance (NMR) and CD spectroscopies. The peptides, 19–21 amino acids in length, were studied in water, except for the C-terminal peptide, which was investigated in DMSO-d_6. NMR showed that all investigated peptides lacked uniquely folded conformations in water at 4 °C and pH 3.0 or at 10 °C and pD 6.6 in DMSO. However, some preferential helix-like conformations for the peptides corresponding to the helices of the folded protein could be seen in solution. These peptides showed characteristic interactions for conformations in both the β- and α-regions of ϕ–ψ space, based on strong $C^{\alpha}H(i)$–$NH(i+1)$ interactions, and on NH–NH, $C^{\alpha}H(i)$–$NH(i+2)$, $C^{\alpha}H(i)$–$NH(i+3)$, and $C^{\alpha}H(i)$–$C^{\beta}H(i+3)$ interactions, respectively. Helical motifs seem not to be the most important factors in determining MHC-binding and/or T-cell recognition. However, there is a tendency that more stabilized secondary structures show higher T-cell stimulation.

The hemoglobin of *Chironomus thummi* (*Chi t* I)[1] has been found to be the main or exclusive allergen of hypersensitivity reactions caused by contact with the Diptera species of this organism (Baur et al., 1982, 1986; Liebers et al., 1993a; Baur & Liebers, 1992). Approximately 20% of exposed people develop hypersensitivity reactions like conjunctivitis, rhinitis, urticaria, and/or bronchial asthma. The *Chi t* I allergy is widespread and serves as a model for human type I hypersensitivity (Tee et al., 1985; Liebers & Baur, 1994). *Chi t* I comprises 12 homologous monomeric or homodimeric protein components of 136–151 amino acids each. *Chi t* I component III consists of 136 amino acids in eight helical and eight nonhelical segments. The three-dimensional structure was first determined by Huber et al. (1970, 1971) at 2.5 Å and was refined by Steigemann and Weber to 1.4 Å (Steigemann & Weber, 1979; Weber et al., 1978).

Extensive immunological studies have been performed on the *Chi t* I proteins, and particularly on its component III. Linear antibody binding regions were defined with monoclonal antibodies as well as with human IgE antibodies. Using monoclonal antibodies and solid-phase bound peptides, three B-cell epitopes could be identified within the regions 1–17, 23–29 (van Kampen et al., 1993), and 91–101 (Mazur et al., 1987, 1988). Sequences 1–17 and 91–101 were shown to recognize human IgE (Mazur et al., 1990) as well. At the T-cell level, nearly the whole sequence was found to stimulate peripheral blood lymphocytes of sensitized individuals. T-cell epitopes of *Chi t* I component III have been identified by using 12 overlapping peptides of 20 amino acids in length each, spanning the entire sequence of component III (Liebers et al., 1994). Differences exist with respect to the percentage of reacting patients and to the degree of the patients' responses (Liebers et al., 1993b).

We present in this paper the determination of solution conformations of peptide fragments derived from *Chi t* I component III using ^{1}H NMR and CD spectroscopies. The following peptides that corresponded to helical and turn segments in the *Chi t* I component III were studied: residues 1–19, 11–30, 31–50, 51–70, 61–80, 71–90, 80–100, 90–110, and 110–130. The same peptides were also used in experiments to assess T-cell stimulation (with the exception of fragments 1–20 and 81–100). The purpose of our study was twofold. First, from the immunological point of view, it was of interest to know whether the peptides still reveal structural motifs of the native antigen and whether there is a relationship between structural features and antigenicity. Since antibodies bind the native antigen, a three-dimensional structure may play a crucial role in antigen recognition. Indeed, conformational epitopes are thought to be the most important determinants of antigenicity on the B-cell level (Laver et al., 1990). It has been observed that peptide immunogens that elicit antibodies that also recognize the cognate sequence in the native protein are often those for which such conformational preferences are observed in aqueous solution (Dyson et al., 1985, 1988a,b, 1992a; Williamson et al., 1986). However, the secondary structure elements may be important for T-cell recognition as well, since T-cells recognize fragments after processing of the native antigen. For example, Rothbard et al. (1989a) suggested that secondary structure elements, such as helices, influence the MHC as well as T-cell binding. According to Margalit (1987), T-cell epitopes are preferentially found in amphipathic helices. A prediction based on this was not able to describe all experimentally defined T-cell epitopes of *Chi t* I component III. The second purpose of our study was to attempt to identify potential folding initiation sites in a protein that is not only α-helical but which also shows large regions

† This work was supported by research grants from the Deutsche Forschungsgemeinschaft.

‡ Max-Planck-Institut für Biochemie.

§ Berufsgenossenschaftliches Forschungsinstitut für Arbeitsmedizin, Bochum.

® Abstract published in *Advance ACS Abstracts,* July 15, 1994.

[1] Abbreviations: *Chi t* I, the hemoglobin of *Chironomus thummi* (nonbiting midges); NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; HOHAHA (TOCSY), homonuclear Hartmann–Hahn spectroscopy (total correlation spectroscopy); 2D, two-dimensional; CD, circular dichroism; TFE, 2,2,2-trifluoroethanol; DMSO-d_6, deuterated dimethyl sulfoxide; MHC, major histocompatibility complex; HLA, human leucocyte antigen; Fmoc, N^{α}-fluorenylmethoxycarbonyl; HPLC, high-performance liquid chromatography.

0006-2960/94/0433-9420$04.50/0

of 3_{10}-helices. The use of peptide fragments of proteins to study initiation steps of protein folding has recently been used in the study of myohemerythrin, a four α-helix bundle protein (Dyson et al., 1992b), plastocyanin, a β-barrel protein (Dyson et al., 1992c), and myoglobin, where two α-helices, one turn region, and a helical hairpin were investigated (Waltho et al., 1993; Shin et al., 1993a,b). Peptides derived from α-helical or turn regions of the protein were often found to show preferences to adopt secondary structure elements like nascent helices or turn-like structures in solution. No preferred secondary structure was found for the β-strand regions of plastocyanin.

MATERIALS AND METHODS

Synthetic Peptides. The peptides corresponding to amino acid sequences 11–30, 31–50, 51–70, 61–80, 71–90, 80–100, 90–110, and 110–130 were assembled by solid-phase synthesis using an automated peptide synthesizer (peptide synthesizer 432A, ABI, Weiterstadt, Germany) and N^{α}-fluorenylmethoxycarbonyl (Fmoc)-protected amino acids (Fields & Noble, 1990). For activation, 2-(1H-benzotriazole-1-yl)-1,1,3,3,-tetramethyluronium hexafluorophosphate (HBTU) was used. The deprotection of the Fmoc-group was carried out in piperidine/dimethylformamide (2:8, v/v). Cleavage of the peptides from the resins was performed in a mixture of trifluoroacetic acid/thioanisole/ethanedithiol (90:5:5, v/v/v) for 2 h at room temperature. The peptides were purified by reverse-phase high-performance liquid chromatography (HPLC) on a silica bead C18 column (size: 20 × 250 mm), and were identified by ion-spray mass spectroscopy on an API III triple-quadrupole mass spectrometer equipped with an IonSpray interface (Sciex, Toronto, ON, Canada). The purity of the peptides was over 95%, as assessed by reverse-phase liquid chromatography. The peptide of fragment 1–19 was a gift from Dr. P. Rücknagel of the Max-Planck-Institute for Biochemistry in Martinsried.

Antigenicity Assay. Antigenicity was measured with the lymphocyte transformation test. Lymphocytes were separated by Ficoll density centrifugation, washed twice with Hanks' balanced salt solution and suspended in RPMI 1640 (Gibco, Eggenstein, Germany) supplemented with 1% L-glutamine (Gibco), 10% human AB serum inactivated by heat (Bayrisches Rotes Kreuz, Munich, Germany), and 1% streptomycin/penicillin (Gibco).

A total of 5×10^4 cells per well were added to 100 μL of allergen in 96-well round-bottom plates. The allergen concentration was 0.32 μmol/L for *Chi t* I and for *Chi t* I component III, and 3.2 μmol/L for the peptides. The control samples contained RPMI instead of the allergen. Cultures were incubated in triplicate at 37 °C in a humidified atmosphere of 5.8% CO_2 for 5 days. Thirty-seven kilobecquerels of tritiated thymidine (NEN, DuPont, Dreieich, Germany) per well were added, and cell proliferation was determined by tritiated thymidine uptake during the last 6 h of the culture period. After the plates were frozen each well was harvested on a cellulose filter, treated with scintillation fluid (Quickszint 1, Zinser Analytic Frankfurt, Germany), and counted with a β-counter (LKB, Turbo, Finland). Reaction of peripheral blood lymphocytes is given as SI (counts per minute with allergen relative to counts per minute without allergen). Background activity was between 500 and 1000 cpm, indicating an SI of 2.0 for 1000–2000 cpm. SIs of responding patients showed a significant difference in mean maximal counts between wells that contained the peptides and wells that contained RPMI at the $p < 0.05$ level (Student's t test).

Sample Preparation of the Peptide Fragments of Chi t I. The NMR samples of the fragments were dissolved in 90% H_2O/10% D_2O at an approximate concentration of 2 mM peptide (10 mM sodium acetate, 4 mM sodium phosphate, 15 mM NaCl) at pH 3.0. The NMR measurements were carried out at 4 °C for all samples except for fragment 110–130, which was not soluble in water at concentrations needed for NMR studies and was therefore investigated in DMSO-d_6 at 10 °C and pD 6.6. Low temperature and low pH were chosen to minimize amide exchange (Wüthrich et al., 1986; Dyson et al., 1988c; Chandrasekhar et al., 1991).

NMR Spectroscopy. NMR spectra were recorded at 500 and 600 MHz on Bruker AM-500 and AMX-600 spectrometers, respectively. All two-dimensional spectra were recorded in the pure-phase absorption mode using time-proportional phase incrementation methods (Marion & Wüthrich, 1983; Ernst et al., 1987). The following spectra were recorded: NOESY (Ernst et al., 1987), with mixing times of 100 and 150 ms, and HOHAHA (TOCSY) (Davis & Bax, 1985), with a MLEV-17 mixing sequence (Bax & Davis, 1985) of 50- and 70-ms duration. For the NOESY spectra, the water resonance was suppressed by the use of a semiselective excitation pulse in which the last 90° pulse in the sequence was replaced by a jump–return sequence with the carrier placed at the position of water (Plateau & Gueron, 1982). For the HOHAHA spectra in H_2O, the water resonance was suppressed by presaturation of the water resonance. For the majority of the spectra, 700 t_1 increments were collected, each with 4K data points, over a spectral width of 10 ppm in both dimensions.

Circular Dichroism Spectroscopy. CD measurements were performed on a Jobin Yvon Dichrograph Mark IV equipped with a thermostated cell holder and connected to a data station for signal averaging and processing. The spectra were recorded at 4 °C with quartz cells of 1 mm optical path length. The peptide concentration was 0.1 mg/mL (approximately 0.05 mM). The spectra were recorded in 30 mM sodium phosphate (the solution was adjusted to pH 3.0 by adding minute amounts of phosphoric acid), and a titration to 10%, 30%, and 60% in TFE was also performed. All data are averages of 10 scans, which are baseline corrected and smoothed.

RESULTS

Solubility and Aggregation. All peptides, with the exception of the C-terminal peptide 110–130, were highly soluble in water. Aggregation phenomena were investigated by studying the concentration dependence of the NMR and CD spectra. For all peptides, additional NOESY spectra were recorded at approximate concentrations of 1 and 0.5 mM with the pHs readjusted to fit the lower concentration of the peptide. No change in line width or chemical shift was noticed upon dilution. Also, the relative magnitudes of the NH(i)–NH($i\pm1$) and $C^{\alpha}H(i)$–NH($i+1$) NOE connectivities were independent of peptide concentration, suggesting that no aggregation phenomena were contributing to the NOEs observed or influencing significantly the conformational preferences of the fragments. The CD study also did not indicate any aggregation of peptides. The CD results are in good agreement with the NMR studies, predicting approximately the same helical content for the peptides at concentrations much lower than those used in the NMR measurements (see below).

Resonance Assignments of Fragments of Chi t I. A sequence-specific resonance assignment was carried out

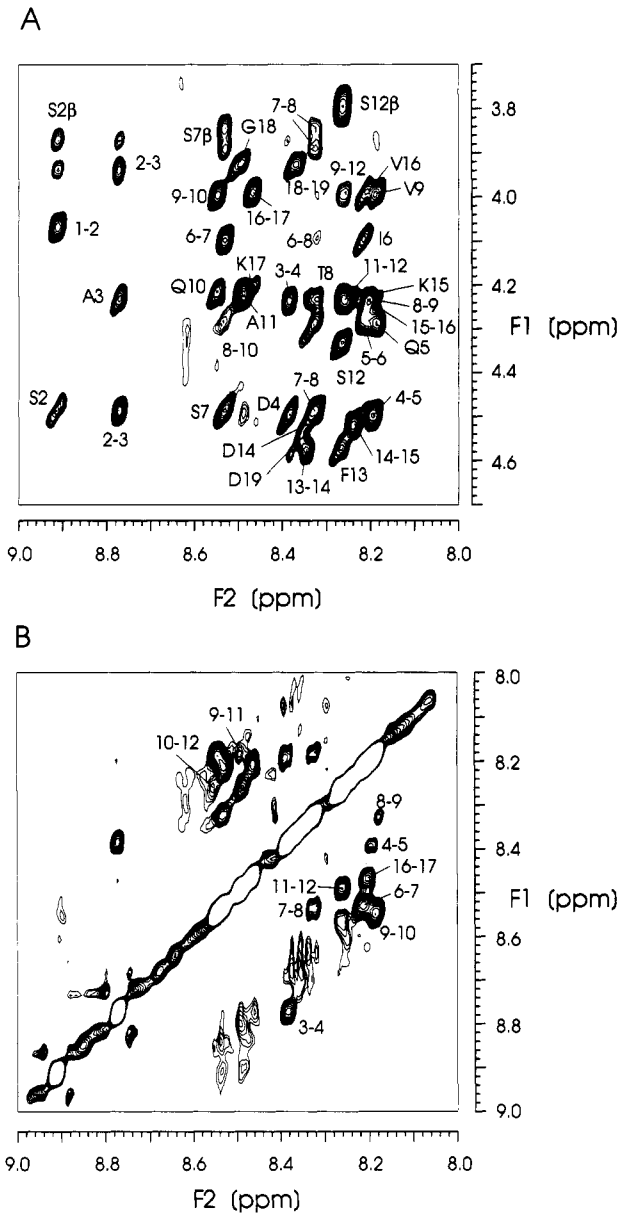

FIGURE 1: NOESY spectrum of fragment 1–19 (pH 3.0) in water at 4 °C with a mixing time of 150 ms. (A) The NH(F_2 axis)–C$^\alpha$H(F_1 axis) region. The cross peaks arise from through-space (<5.0 Å) connectivities. Intraresidue cross peaks are labeled with one-letter abbreviations and interresidue connectivities are indicated by simple numbers. (B) The NH(F_1 axis)–NH(F_2 axis) region of the NOESY spectrum.

according to the procedure of Wüthrich (1986). HOHAHA spectra recorded with 50- and 70-ms mixing times were used to identify amino acid spin systems by successively detecting direct, single, and relayed through-bond connectivities. NOESY spectra were used to observe through-space (<5 Å) connectivities. As an example, the NOESY spectrum of the fragment 1–19 at 4 °C and pH 3.0 is shown in Figure 1. Lists of the NMR assignments for all fragments are given in Tables I–X of the supplementary material, including the TFE study of fragment 51–70.

Determination of Preferential Secondary Structure Elements from NMR. A summary of the diagnostic NOEs for all investigated peptide fragments is shown in Figure 2a–i. The NMR data for all of the fragments showed a lack of long-range NOEs (between amino acids further apart than four residues). Ordered conformations in peptides give rise to distinctive patterns of diagnostic NOE connectivities. From such NOEs it is frequently possible in a qualitative manner to assess the populations of different types of conformations (Dyson et al., 1988; Bruch et al., 1989). An extended strand gives rise to very strong C$^\alpha$H(i)–NH(i+1) NOEs and no NH–NH intensities for each pair of residues. This is in contrast to the α-conformation, which involves strong NH(i)–NH($i\pm1$) NOEs accompanied by weak C$^\alpha$H(i)–NH(i+1) intensities (Wüthrich, 1986). Helical structures, loose helical structures, and turns can be also detected via the presence of characteristic medium range interactions: the C$^\alpha$H(i)–NH(i+2), C$^\alpha$H(i)–NH(i+3), NH(i)–NH(i+2), and C$^\alpha$H(i)–C$^\beta$H(i+3) connectivities. The measured NOEs are average values resulting from a mixture of extended conformations and α-conformations. For any pair of residues, mostly strong NOEs of each state will contribute to the average NOE. The presence of strong NH–NH connectivities indicates that some fraction of the population of the peptide occupies an α-region of ϕ,ψ space. Because the unfolded conformers give rise to strong C$^\alpha$H(i)–NH(i+1) NOEs for each pair of residues in the peptide, obscuring the NOEs originating from secondary structure elements that contain weak C$^\alpha$H(i)–NH(i+1) connectivities, the C$^\alpha$H(i)–NH(i+1) NOEs cannot be used for characterization of helices and turns in peptides. An approximate estimate of the helical or nascent helical content of *Chi t* I can be assessed from measurement of the intensities of the C$^\alpha$H(i)–NH(i+1) and NH–NH NOEs (Czisch et al., 1993; Waltho et al., 1993). Assuming that both random conformations and helical conformations contribute to the intensities of the C$^\alpha$H(i)–NH(i+1) cross-peaks and that all helical conformations (nascent as well as more stabilized helical conformations) contribute exclusively to the NH–NH cross-peaks. The ratio of the peaks intensities I

$$R = I[\mathrm{NH}(i)\text{–}\mathrm{NH}(i+1)]/I[\mathrm{C}^\alpha\mathrm{H}(i)\text{–}\mathrm{NH}(i+1)]$$

gives directly the relative population in the helical conformation at residue i. If the mean value of all these ratios over the whole peptide is calculated, the result is an overall helicity, which is equivalent to the helicity measured by CD spectroscopy.

Secondary Structure from CD. Secondary structure determination by CD is based on the characteristic ellipticity of helices, β-sheets, and random coil. Ordered α-helical structures of at least two or three turns can be identified in CD spectra by characteristic minima at 208 and 222 nm. Small, loose, and nascent helices on the other hand do not give exactly these minima (Manning et al., 1988). Turn-like structures are normally difficult to detect in CD. The presence of these loose conformations and turns make exact calculations of helical content impossible (Provencher & Glöckner, 1981). For evaluation of the peptide conformations, we followed the procedure described by Greenfield and Fasman (1969). The percentages of helical conformation from the CD measurements were estimated from the location and intensities of the CD minima.

Conformational Preferences for Fragment 1–19. This peptide covers the whole first helix of the intact protein. The C$^\alpha$H–NH and the NH–NH regions of the NOESY spectrum collected with a mixing time of 150 ms are shown in Figure 1. Besides strong C$^\alpha$H(i)–NH(i+1) cross peaks, the NMR spectrum shows strong NH–NH contacts and also six diagnostic C$^\alpha$H(i)–C$^\beta$H(i+3) interactions, characteristic for a helical motif, starting with the Ser 2–Gln 5 contact (Figure 2a). The missing NH–NH contact between Ser 2 and Ala 3 indicates that the helix starts at position 3. It ends at Asp 14. The first three to four residues of this helix show the most intense NH–NH contacts, compared to the corresponding

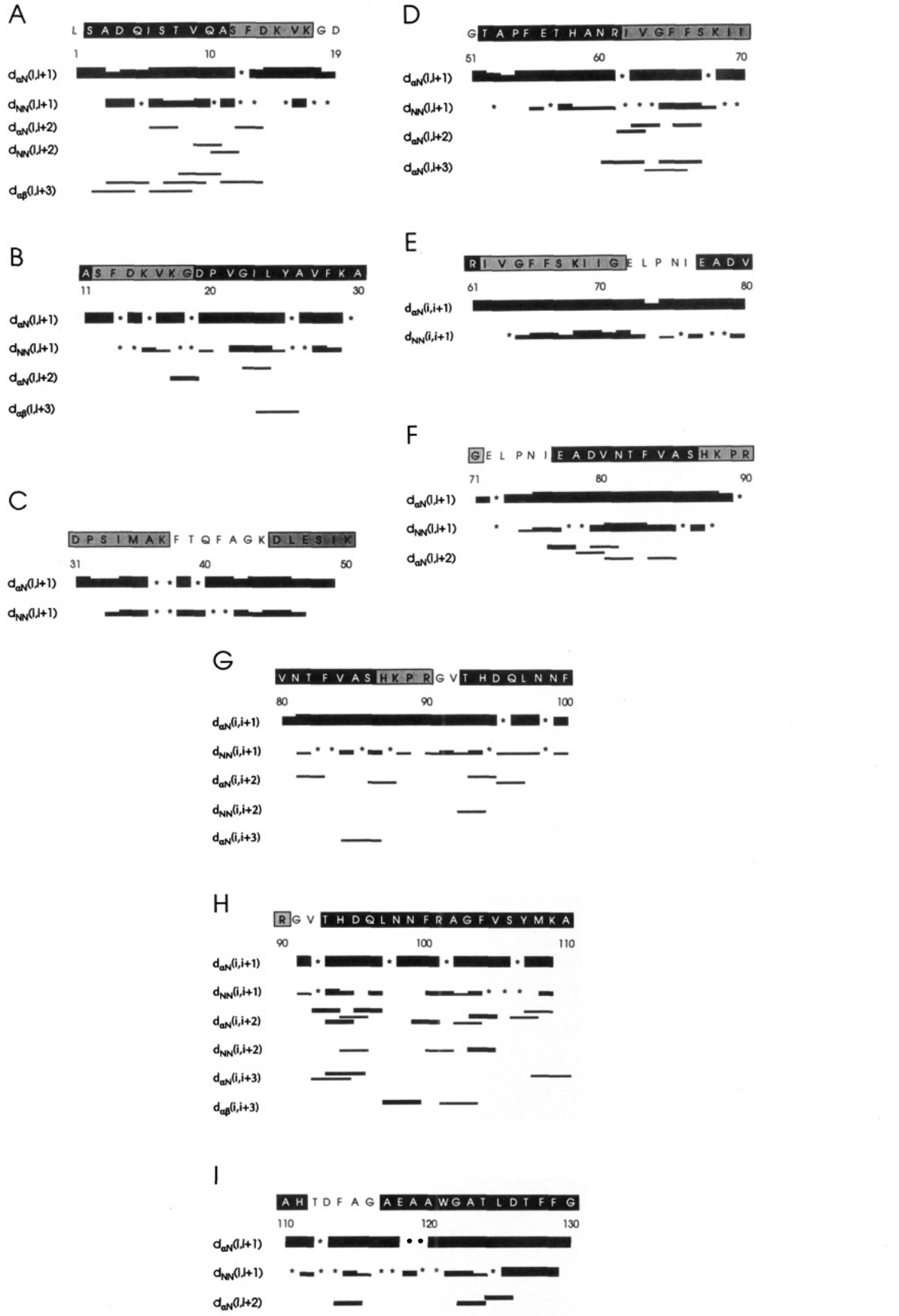

FIGURE 2: Summary of the diagnostic NOEs. The intensities of the NOEs are reflected in the heights of the bars. The $C^{\alpha}H(i)$–$C^{\delta}H(i+1)$(Pro) and the NH(i)–$C^{\delta}H(i+1)$(Pro) NOEs are shown along the same line as the $C^{\alpha}H(i)$–NH(i+1) and the NH(i)–NH(i+1) NOEs, respectively. For $C^{\alpha}H(i)$–NH(i+1) and NH(i)-NH(i+1) NOEs, an asterisk indicates the data that could not be positively determined because of spectral overlap. Blank spaces indicate the lack of a positively established NOE. For medium-range NOEs, only positively determined peaks are shown. In the top entry, the primary structure of the corresponding fragment is given. The type of shading indicates the type of secondary structure found in the intact protein. Black areas correspond to α-helix, and dark shading corresponds to 3_{10}-helix. (A) Fragment 1–19. (B) Fragment 11–30. (C) Fragment 31–50. (D) Fragment 51–70. (E) Fragment 61–80. (F) Fragment 71–90. (G) Fragment 80–100. (H) Fragment 90–110. (I) Fragment 110–130.

Table 1: Summary of the Overall Helical Content in the Peptides[a]

peptide	NMR	CD
1–19	15%	15% (70%)
11–30	<10%	<10% (40%)
31–50	≤10%	<10% (40%)
51–70	10–15% (60%)	≤10% (50%)
61–80	10–15%	<10% (25%)
71–90	10%	15% (70%)
80–100	<10%	<10% (30%)
90–110	10–15%	15% (60%)
110–130	≤10%	not measured

[a] For the NMR data, the relative population in a helical conformation was calculated for each residue, as described in the text, and averaged over all residues. CD results were obtained as described in the text. Data in parentheses refer to the measurements performed in 60% TFE.

$C^{\alpha}H(i)$–$NH(i+1)$ peaks, indicating a relative population of ~50% in helical structure for these residues. This amount becomes ~15% for other residues in the peptide up to Asp 14. At the C-terminus of the peptide, there is only one unambiguous NH–NH contact between Val 16 and Lys 17. No $C^{\alpha}H(i)$–$C^{\beta}H(i+3)$ connectivities were seen for this part of the sequence. The presence of two $C^{\alpha}H(i)$–$NH(i+2)$ interactions indicates that there is also a small population of nascent helix or, recalling that the C-terminal part adopts 3_{10}-helicity in the intact protein, probably the beginning of a 3_{10}-helix. If the remaining residues are approximately 5% helical, an overall helicity for the peptide is ~15% (Table 1). The NMR results were confirmed by CD, where a helical population of approximately 15% is found in the pure water. This value increases with increasing amount of added TFE to a final value of roughly 70% helix for a solution containing 60% TFE. Under these conditions, the whole peptide contributes to the secondary structure, except for three or four residues at the N and C ends.

Conformational Preferences for Fragment 11–30. In the protein, this region covers the C-terminus of the 3_{10}-helix A from Ala 11 to Asp 19 and the complete α-helix B from Pro 20 to Ala 30 (Figure 2b). The synthetic peptide showed weaker NH–NH contacts for residues Val 16 and Lys 17 than the corresponding NH–NH contacts in peptide 1–19. Indications for a nascent helix from Lys 15 to Gly 18, and perhaps even two residues longer at the N-terminus, were also detected. There is an additional helical motif starting with Val 21 and ending with Lys 29, as indicated by a sequence of NH–NH contacts. Two additional cross peaks were detected: one $C^{\alpha}H(i)$–$C^{\beta}H(i+3)$ contact between Ile 23 and Ala 26, and one $C^{\alpha}H(i)$–$NH(i+2)$ contact between Gly 22 and Leu 24. Other diagnostic NOEs could not be found, mainly due to overlap with the diagonal and other cross peaks. Thus, the NOE pattern indicates a nascent helix from 21 to 29, which is stabilized in the middle, connected to a weaker nascent helix from residues at the N-terminus. The overall helicity calculated from the NMR data is given in Table 1.

Conformational Preferences for Fragment 31–50. This region covers the complete 3_{10}-helix C from residue 31 to residue 37 in native hemoglobin. Phe 38 to Lys 44 is a nonhelical segment, which is followed by another complete 3_{10}-helix from residue Asp 45 to Lys 50 (Figure 2c). In the corresponding peptide, strong sequential $C^{\alpha}H(i)$–$NH(i+1)$ contacts and 10 sequential NH–NH cross peaks were observed. Together with four possible NH–NH contacts, which are obscured by diagonal or other peaks, these NOEs span approximately 70% of fragment 31–50. Only two residues at the N-terminus (including Pro 32) and three residues at the C-terminus do not show these types of peaks. However, no diagnostic medium-range connectivities could be found unambiguously. The peptide builds a 14-residue-long nascent helix, perhaps distorted in the middle.

FIGURE 3: Summary of the diagnostic NOEs for 51–70 with 60% TFE/40% H_2O.

Conformational Preferences for Fragment 51–70. The X-ray structure of the protein shows a regular helix (helix E) from Thr 52 to Arg 61. Helix E proceeds to a 3_{10}-helix (Val 63–Gly 71) that is linked to the heme group. The corresponding peptide shows also many indications for a helical structure (Figure 2d). Nine unambiguous sequential NH–NH contacts were detected; seven more possible contacts were obscured by other peaks or by the diagonal peaks. These contacts start directly after proline 56 and cover the rest of the peptide. In addition, three unambiguous $C^{\alpha}H(i)$–$NH(i+2)$ contacts and three $C^{\alpha}H(i)$–$NH(i+3)$ peaks were found. No diagnostic $C^{\alpha}H(i)$–$C^{\beta}H(i+3)$ or $C^{\alpha}H(i)$–$NH(i+4)$ peaks were detected. There are no apparent conformational preferences for the N-terminal part of this fragment (Phe 55–Ala 59). The $C^{\alpha}H(i)$–$NH(i+3)$ peaks found for residues N60-V63, V63-F66, and G64-S67 indicate that a certain population of the peptide has a more stabilized helical conformation. Because of the missing $C^{\alpha}H(i)$–$C^{\beta}H(i+3)$ peaks, we believe that this part preferably forms a 3_{10}-helix [with only weak $C^{\alpha}H(i)$–$C^{\beta}H(i+3)$ contacts, which were not detectable in our spectra] instead of an α-helix [where these peaks typically are clearly visible and even more intense than the $C^{\alpha}H(i)$–$NH(i+3)$ peaks (Waltho et al., 1993)]. Also, no $C^{\alpha}H(i)$–$NH(i+4)$ peaks characteristic for an α-helix were found. Again, the first two residues at the N-terminus and the last two residues at the C-terminus do not participate in the helix.

In addition, this peptide was also investigated by NMR in the presence of 60% TFE. The pH of the sample was 3.8; all other parameters were the same as in the other NMR spectra. The diagnostic NOE connectivities are shown in Figure 3. It is clearly visible that a large part of this fragment shows all connectivities characteristic of an α-helix. This helix starts at Glu 56, perhaps at Phe 55, and ends at Ile 70, with the best defined region located in the region of the 3_{10}-helix found in pure water as described above. The first four N-terminal residues do not show clear helical behavior; nevertheless, there are two NH–Pro(δ-CH_2) contacts visible (Ala 53–Pro 54–Phe 55), indicating some ordered secondary structure.

Conformational Preferences for Fragment 61–80. Intact hemoglobin shows two helical parts, starting with a regular type 3_{10}-helix E from Val 63 to Gly 71 (Figure 2e). At Ile 62, there is a loop connecting to the heme group. From Glu 72 to Ile 76, there is an open turn; this is followed by the

N-terminus of α-helix F. In the corresponding peptide, many NH–NH contacts were found, starting at Gly 64 or even at Val 63, and ending at Val 80, with an interruption between Leu 73 and Pro 74. No medium-range contacts could be positively established. This indicates that there is a loose nascent helix starting at Gly 64, with some disruption around the proline. The N-terminal part of this nascent helix seems to be better defined, as indicated by the increased magnitudes of the NH–NH contacts.

Conformational Preferences for Fragment 71–90. The first residue of this fragment belongs to helix E in the intact protein; the rest comprises α-helix F (Glu 77–Ser 86) and a distorted 3_{10}-helix between His 87 and Arg 90. An open turn was found between Glu 72 and Ile 76 in the intact protein. The corresponding peptide (Figure 2f) showed 10 sequential NH–NH contacts and five $C^{\alpha}H(i)$–$NH(i+2)$ contacts that were located in the middle of the peptide at the position of the F α-helical part. No medium-range NOEs were found before Pro 74 and after His 87. The peptide seems to adopt a nascent helix that is most stable at the center of the peptide.

Conformational Preferences for Fragment 80–100. Fragment 80–100 covers part of α-helix F of the hemoglobin (up to Ser 86), followed by a distorted 3_{10}-helix (His 87–Gly 91) and an extended conformation (Gly 91 and Val 92). The last seven residues belong to α-helix G. The NMR spectra of this fragment indicate that the whole peptide has a tendency to adopt nascent helical structure, including the gap between helices F and G (Figure 2g). In addition to 12 NH–NH contacts, six medium-range connectivities could be found: four $C^{\alpha}H(i)$–$NH(i+2)$ contacts and the $C^{\alpha}H(i)$–$NH(i+3)$ and $NH(i)$–$NH(i+2)$ contacts.

Conformational Preferences for Fragment 90–110. In the intact protein, the region from 90 to 110 starts with an extended conformation (Gly–Val) and is followed by nearly the complete α-helix G (with the terminal His 111 missing; Figure 2h). In the synthetic fragment, many medium-range connectivities were detected in addition to nine unambiguous NH–NH contacts. This indicated a nascent helix covering the complete helix G, with a tendency to adopt an α-helical conformation in the middle of this region [as indicated by the $C^{\alpha}H(i)$–$C^{\beta}H(i+3)$ contacts between Leu 97–Phe 100 and Arg 101–Phe 104]. Three $C^{\alpha}H(i)$–$NH(i+3)$ interactions were visible at the beginning and at the end of the G helix. Together with the missing $C^{\alpha}H(i)$–$C^{\beta}H(i+3)$ contacts, these might indicate that the nascent helix adopts the characteristic features of a 3_{10}-helix.

Conformational Preferences for Fragment 110–130. This fragment was the only peptide not soluble in water at concentrations necessary for NMR measurements. We therefore investigated it in DMSO-d_6 at 10 °C and pD 6.5. In the intact protein, this region contains the end of α-helix G (Ala 110–His 111), a nonhelical region of five residues, and the complete α-helix H (Ala 117–Gly 130). The peptide showed 11 sequential NH–NH contacts and three $C^{\alpha}H(i)$–$NH(i+2)$ contacts (Figure 2i). Two of the later were in the region of helix H, where the most intense NH–NH contacts were also found. We assume that there is a nascent helix for this segment. The rest of the peptide does not show any clear-cut patterns. There were some cross peaks in the nonhelical fragment, indicating that this region might also adopt a folded conformation when not in the intact protein.

DISCUSSION

Since the formation of secondary structure is thought to be the earliest event in protein folding, linear peptides are often used as models for initiation sites of folding (Wright et al., 1988). In this model, secondary structure elements develop first and are in rapid dynamic equilibrium with unfolded conformations of the protein. The secondary structure is stabilized by short- and medium-range interactions. In later folding steps, the polypeptide chain collapses and long-range interactions start to act. In linear model peptides, only the local amino acid interactions influence formation and stabilization of the structure. It is clear from our results that elements of secondary structure, in the form of nascent helices and more ordered helical conformations, are abundant in short peptide fragments of the hemoglobin of *C. thummi*. In most cases, these secondary structure elements were found to be in the same locations as helices in the intact protein, whereas the weakest secondary structure elements were found in parts of the hemoglobin which show no regular secondary structure (Figure 4). The regions that form 3_{10}-helices in the intact protein also form helical conformations in the peptides. The stabilities of these conformations were found to be indistinguishable from those of peptides from α-helical parts of the protein. Our results thus complement the data of a study on peptide fragments of an all-α-helical protein, myohemerythrin (Dyson et al., 1992b).

Figure 4 shows a comparison of the synthetic peptide and protein structures. The primary structure of intact hemoglobin *Chi t* I component III is shown in the top entry together with the secondary structure elements as found by X-ray crystallography. In comparison to the secondary structure of the intact protein, the α-helical motifs in peptide 1–19 are less well defined but are still clearly present in the peptide. The 3_{10}-helix from 12 to 19 is only visible as a nascent helix. The N- and C-terminal ends of the peptide are in a nonhelical conformation. This fraying of the termini of linear peptides has been already observed by Dyson et al. (1992b). Helix B of the protein is present in peptide 11–30 as a nascent helix at the same location. Many hydrophobic residues are located here that may account for the increased stability of the structure (Richardson & Richardson, 1988). For the N-terminus of the peptide, the helical motif of helix A in the intact hemoglobin is conserved as a nascent helix. In the region of helices C and D, peptide 31–50 exhibits a nascent helix, possibly interrupted in the middle by a turn-like structure in the protein nonhelical segment. We found contacts between residues in the nonhelical segment that should not be visible if the structure is exactly like in the intact protein. The N-terminal part of helix E, which is very regular in the protein, seems to be more loose in the peptide corresponding to region 51–70. Probably due to the missing heme group in the peptide, the second half of the helix (originally type 3_{10}) starts earlier. Surprisingly, in the peptide this helix is better defined than the first loose helix. This might be due to the fact that this region is rich in hydrophobic residues. Compared to the intact protein, fragment 61–80 showed slightly different secondary structure motifs. The first 3_{10}-helix E is only visible at the N-terminus as a loose nascent helix. Also, cross peaks characteristic for a nascent helix were detected in the region of the open turn and at the beginning of helix F. These differences might be explained by the missing heme group in the synthetic peptide. Fragment 71–90 shows nascent helix in the middle of helix F. The region between helices F and G of the hemoglobin is covered by peptide 80–100, where a nascent helix was found. Peptide 90–110 showed a tendency to fold into a helical conformation in the same region as helix G in the protein. Again, a region with many hydrophobic residues is located in this part. The N-terminal peptide 110–

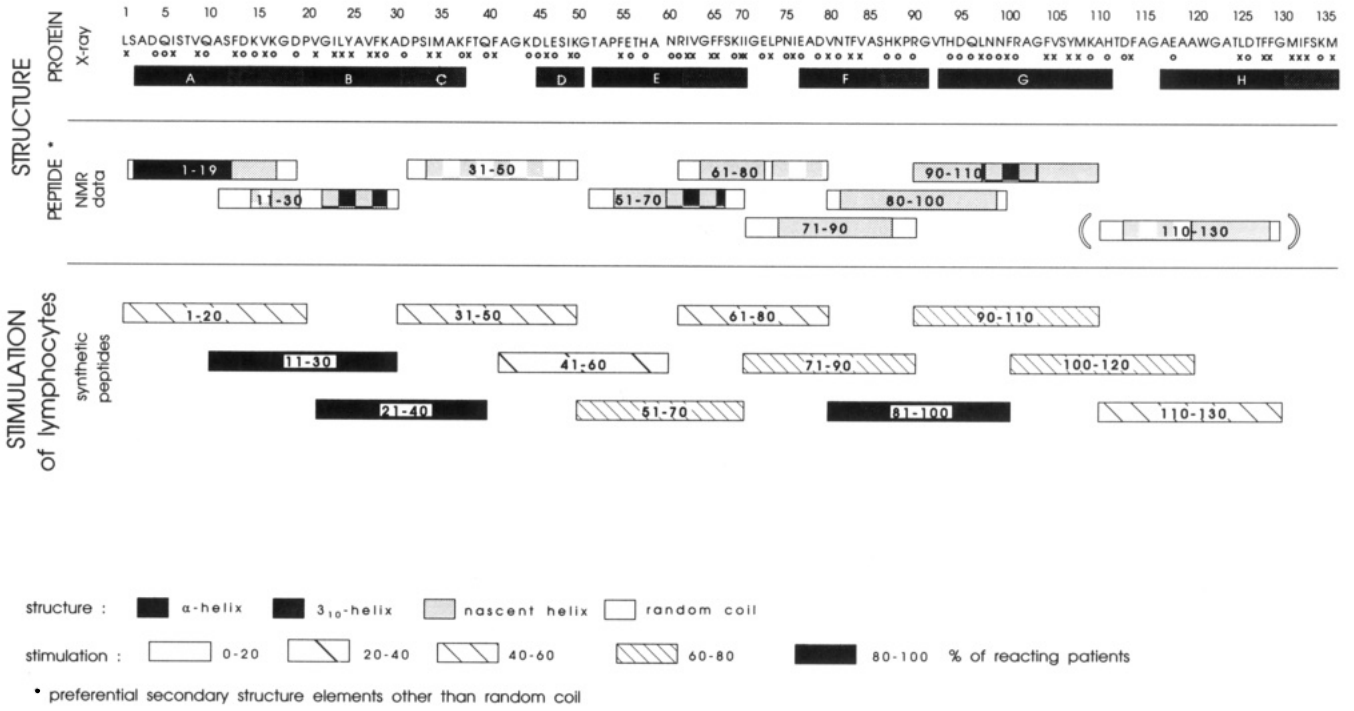

FIGURE 4: Summary of the preferred secondary structure elements and comparison with T-cell stimulation. In the top entry, the primary structure of *Chi t* I component III is given. Hydrophobic residues are marked with "x", hydrophilic residues with "o". The secondary structure elements found in the X-ray structure for the intact protein, and the preferred structure elements in the fragments as detected by NMR are indicated as described in the figure. Regions with transitions between different secondary structure motifs are indicated by hatching. The results on the stimulation of lymphocytes are shown in the last entry, with the percentage of reacting patients indicated in the figure [for details, see Liebers et al. (1993b)].

130 was investigated in DMSO instead of water. We found that helix H was still conserved, in the form of a nascent helix; also, some structure between helices G and H was found.

In conclusion, we showed that all peptides examined revealed structural elements similar to those of the native antigen. The best defined helices were found in fragments 1–19 and 90–110, which both contain antibody-binding determinants. These findings fit well with the investigation of Dyson et al. (1985), which showed that immunodominant sites show high preferences to adopt folded conformations; however, at the B-cell level, the correlation between secondary structure elements and antigenicity is not so strong. For the HIV 1 protein p18, it was shown that larger peptides are more antigenic than smaller ones (Marbrouk et al., 1993). This could not be explained by the fact that larger fragments have higher propensities to adopt a particular secondary structure. As measured by circular dichroism, random coil conformations predominated in all peptides in aqueous solutions, and α-helices were found in short as well as in large peptides (Marbrouk et al., 1993). For T-cells that are stimulated by antigen-derived fragments bound to MHC molecules, amphipathic helices were reported to behave as epitopes (Margalit et al., 1987; Rothbard et al., 1989a). There is some evidence that helices are not a prerequisite for HLA-binding (Grey et al., 1989). In spite of this, secondary structural features of the epitope, such as helices, may lead to better recognition by the T-cell receptor (Rosloniec et al., 1993; Rothbard et al., 1989b, 1991). The NMR measurements thus show that the peptides used for epitope mapping *in vitro* correspond to the native antigen in terms of their structural features.

At the T-cell level, all investigated peptides showed antigenicity (Liebers et al., 1993b) (Figure 4). Peptide 11–30 stimulated lymphocytes in nine out of 11 tested patients (82%). However, although fragment 1–19 revealed a much more stabilized helix, the corresponding peptide used in clinical trials (residues 1–20) stimulated only six out of a total of ten patients (60%). A similar situation was found for fragments 80–100 and 90–100. Fragment 80–100 showed stimulation of 10 patients from a total of 12, but the neighboring sequence 90–110, which had a greater tendency for a stabilized structure in our study, only stimulated 7 out of 12 tested patients. These findings suggest that helical motifs may not be the most important factors in determining MHC-binding and/or T-cell recognition. Nonetheless, the other fragments that are only weakly structured (31–50 and 61–80) showed lymphocyte stimulation of less than 60% of the patients, whereas the better defined fragments (51–70, 71–90, and 90–110) stimulated more than 60% of the patients.

ACKNOWLEDGMENT

We thank P. Rücknagel for the preparation of fragment 1–19, Ms. E. Weyher-Stingl for assistance with the circular dichroism measurements, and Michael Becker for discussions.

SUPPLEMENTARY MATERIAL AVAILABLE

Tables of the chemical shifts of fragments between residues 1–19, 11–30, 31–50, 51–70, 61–80, 71–90, 80–100, and 90–110 in water and a table of the chemical shifts of fragment 51–70 measured in 60% TFE/40% water (20 pages). Ordering information is given on any current masthead page.

REFERENCES

Baur, X., & Liebers, V. (1992) *Int. Arch. Occup. Environ. Health 64*, 185–188.

Baur, X., Dewair, M., Fruhmann, G., Aschauer, H., Pfletschinger, J., & Braunitzer, G. (1982) *J. Allergy Clin. Immunol. 69*, 66–76.

Baur, X., Aschauer, H., Mazur, G., Dewair, M., Prelicz, H., & Steigemann, W. (1986) *Science 233*, 351–354.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–366.

Bruch, M. D., McKnight, C. J., & Gierasch, L. M. (1989) *Biochemistry 28*, 8554–8561.
Chandrasekhar, K., Profy, A. T., & Dyson, H. J. (1991) *Biochemistry 30*, 9187–9194.
Czisch, M., Schleicher, M., Hörger, S., Voelter, W., & Holak, T. A. (1993) *Eur. J. Biochem. 218*, 335–344.
Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820–2821.
Dyson, H. J., Cross, K. J., Hougten, R. A., Wilson, I. A., Wright, P. E., & Lerner, R. A. (1985) *Nature 318*, 480–483.
Dyson, H. J., Rance, M., Hougten, R. A., Lerner, R. A., & Wright, P. E. (1988a) *J. Mol. Biol. 201*, 161–200.
Dyson, H. J., Rance, M., Hougten, R. A., Wright, P. E., & Lerner, R. A. (1988b) *J. Mol. Biol. 201*, 201–217.
Dyson, H. J., Lerner, R. A., & Wright, P. E. (1988c) *Annu. Rev. Biophys. Biophys. Chem. 17*, 305–324.
Dyson, H. J., Norrby, E., Hoey, K., Parks, D. E., Lerner, R. A., & Wright, P. E. (1992a) *Biochemistry 31*, 1458–1463.
Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A., & Wright, P. E. (1992b) *J. Mol. Biol. 226*, 795–817.
Dyson, H. J., Sayre, J. R., Merutka, G., Shin, H. C., Lerner, R. A., & Wright, P. E. (1992c) *J. Mol. Biol. 226*, 819–835.
Ernst, R. R., Bodenhausen G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.
Fields, G. B., & Noble, R. L. (1990) *Int. J. Pept. Protein Res. 35*, 161–214.
Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 8*, 4108–4116.
Grey, H. M., Demotz, S., Buus, S., & Sette, A. (1989) *Cold Spring Habor Symp. Quant. Biol. 54*, 393–399.
Huber, R., Epp, O., & Formanek, H. (1970) *J. Mol. Biol. 52*, 349–354.
Huber, R., Epp, O., Steigemann, W., & Formanek, H. (1971) *Eur. J. Biochem. 19*, 42–50.
Kampen, van V., Chen, Z., Raulf, M., Czuppon, A., & Baur, X. (1993) *Allergo J. 2*, 79/21–83/25.
Laver, W., Air, G., Webster, R., & Smith-Gill, S. (1990) *Cell 61*, 553–556.
Liebers, V., & Baur, X. (1994) *Clin. Exp. Allergy* (in press).
Liebers, V., Hoernstein, M., & Baur, X. (1993a) *Allergy 48*, 236–239.
Liebers, V., Raulf, M., Mazur, G., Modrow, S., Baur, X. (1993b) *J. Allergy Clin. Immunol. 92*, 334–339.
Mabrouk, K., Moulard, M., Gluckman, J., Romi, R., Rochat, H., Rietschoten, J., & Bahrout, E. (1993) *Mol. Immunol. 30*, 503–512.
Manning, M. C., Illangasekare, M., & Woody, R. W. (1988) *Biophys. Chem. 31*, 77–86.
Margalit, H., Spouge, J., Cornette, J., Cease, K., Delisi, C., & Berzofsky, J. (1987) *J. Immunol. 138*, 2213–2219.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Mazur, G., Becker, W. M., & Baur, X. (1987) *J. Allergy Clin. Immunol. 80*, 876–883.
Mazur, G., Baur, X., Modrow, S., & Becker, W. M. (1988) *Mol. Immunol. 25*, 1005–1010.
Mazur, G., Baur, X., & Liebers, V. (1990) *Monogr. Allergy 28*, 121–137.
Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.
Provencher, S. W., & Glöckner, J. (1981) *Biochemistry 20*, 33–37.
Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652.
Rosloniec, E., Beard, S., & Freed, J. (1993) *Mol. Immunol. 30*, 491–501.
Rothbard, J., & Gefter, M. (1991) *Annu. Rev. Immunol. 9*, 527–565.
Rothbard, J., Busch, R., Hill, C., & Lamb, J. (1989a) *Cold Spring Habor Symp. Quant. Biol. 54*, 431–443.
Rothbard, J., Busch, R., Howland, K., Bal, V., Fenton, C., Taylor, W., & Lamb, J. (1989b) *Int. Immunol. 1*, 487–495.
Shin, H. C., Merutka, G., Waltho, J. P., Wright, P. E., & Dyson, H. J. (1993a) *Biochemistry 32*, 6348–6355.
Shin, H. C., Merutka, G., Waltho, J. P., Dyson, H. J., & Wright, P. E. (1993b) *Biochemistry 32*, 6356–6364.
Steigemann, W., & Weber, E. (1979) *J. Mol. Biol. 127*, 309–338.
Tee, R., Crauston, P., Dewair, M., Prelicz, H., Baur, X., & Kay, A. (1985) *Clin. Allergy 15*, 335–343.
Waltho, J. P., Feher, V. A., Merutka, G., Dyson, H. J., & Wright, P. E. (1993) *Biochemistry 32*, 6337–6347.
Weber, E., Steigeman N., Jones, T. A., & Huber, R. (1978) *J. Mol. Biol. 120*, 327–336.
Williamson, M. P., Hall, M. J., & Handa, B. K. (1986) *Eur. J. Biochem. 158*, 527–536.
Wright, P. E., Dyson, H. J., & Lerner, R. A. (1988) *Biochemistry 27*, 7167–7175.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley, New York.