THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 272, No. 1, Issue of January 3, pp. 362–368, 1997
© 1997 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in U.S.A.

NMR Studies on the Flexibility of the Poliovirus C3 Linear Epitope Inserted into Different Sites of the Maltose-binding Protein*

(Received for publication, May 17, 1996, and in revised form, September 13, 1996)

Anne Lecroisey‡§, Pierre Martineau¶, Maurice Hofnung¶, and Muriel Delepierre‡

From the ‡Laboratoire de Résonance Magnétique Nucléaire, CNRS URA 1129 and the ¶Unité de Programmation Moleculaire et Toxicologie Génétique, CNRS URA 1444, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

A set of permissive positions that tolerate insertions/deletions without major deleterious consequences for the binding activity of the protein was previously identified in the maltose-binding protein. The C3 epitope from poliovirus VP1 protein (^{93}DNPASTTNKDK^{103}) was inserted into eight of these positions and two nonpermissive control sites. NMR studies were performed on the MalE protein, the insertion/deletion mutants, and the C3MalE hybrids to selectively determine the flexible regions in these proteins. Comparison of the C3 epitope mobility in the different hybrid proteins indicates that, whatever its insertion site and independently from the specific sequences of its linkers, the epitope is mostly flexible. The vector protein was shown to unfold partially only in the two C3MalE hybrids that correspond to nonpermissive positions. For one of them (insertion at site 339), both sides of the insert are flexible, and at most one side for all the other hybrids. This result correlates with the antigenicity data on the inserted epitope (Martineau, P., Leclerc, C., and Hofnung, M. (1997) *Mol. Immunol,* in press.

Natural or synthetic peptide fragments of proteins are widely used as antigenic and immunogenic probes. However, as a rule, anti-peptide antibodies show less affinity for the native protein than antibodies raised against the protein itself, and similarly, antibodies raised against the native protein bind antigenic peptides less tightly than the native protein. Linear peptides generally poorly mimic the conformation of the antigenic determinants because of their largely unfolded structure in aqueous solution, and restriction of their conformational space has been used for decades to obtain a better mimicry (1). There are a great variety of approaches to carry out this restriction; however, most of them rely on the knowledge of the structure of the epitope in its natural environment: for instance, epitopes in loops will often be mimicked by cyclic peptides, whereas stabilization of elements of secondary or tertiary structure will generally be performed by an appropriate modification of the original sequence (2).

Expression of peptides as fusions within a carrier protein, the maltose-binding protein (MalE), was proposed by Hofnung *et al.* (3). This procedure allows anti-peptide antibodies to be elicited without peptide synthesis and without knowledge of the epitope conformation. A set of "permissive" sites was determined in the MalE protein. These sites are regions of the protein that accept modifications of their primary sequence without major deleterious consequences for the biological activity, structure, and cellular localization of the hybrids. Random insertion of a *Bam*HI octanucleotide linker into the *malE* gene was first described to generate series of insertion/deletion mutants (4, 5). Depending on the reading frame and surrounding nucleotides, there was a linker insertion that contained 2 or 3 amino acids and was accompanied in most cases by small deletions corresponding to the loss of a few residues. The permissive sites were mainly located in structured regions on the surface of MalE (6). They were shown to accept the insertions of heterologous peptides of various length and origin (7–11). One of the tolerated insertions was the C3 epitope from poliovirus type I (Mahoney strain). C3 is a neutralizing epitope (12). In the virus, it forms a hydrophilic loop comprised between two antiparallel β-strands and encompassing residues 93–103 (DNPASTTNKDK) at the surface of the VP1 envelope protein (13).

Peptide DNPASTTNKDK was inserted into 11 different sites in the MalE protein; its antigenic properties (recognition by antibodies) and immunogenic properties (ability to elicit an immune response) were studied for 10 of these sites (eight permissive and two nonpermissive) (11).[1] Depending on the insertion site, the C3 peptide was flanked by linkers of 2 or 3 amino acids with specific sequences (see Table II). It is likely that the fusions at both ends may limit the conformational space of a peptide inserted into a carrier protein in comparison with the free peptide. Its intrinsic flexibility remains constant, but its flexibility, when inserted into the vector, may depend on the site of insertion. It was therefore tempting to investigate whether the flexibility of the C3 epitope was dependent on its localization in the hybrid proteins.

NMR is currently used to study the three-dimensional structure of proteins up to a M_r of 20,000. Studies of larger proteins are hindered by the loss of spectral resolution. First, there is an overlap of the individual resonances from the increased number of protons. Second, there is a severe broadening of these resonances due to the slow overall rotational tumbling of the protein associated with short transverse relaxation time (T_2) values. Line widths predicted by a simple model for macromolecules are indeed in the range of 50–100 Hz or even greater for motions on the order of 5×10^{-8} to 5×10^{-6} s (14). Since in a two-dimensional experiment, the cross-peak intensities decrease as T_2, most of the signals merge into the background. If, however, some regions of the molecule present a local flexibility, that is when internal motion occurs on a much faster time scale, on the order of 10^{-9} to 10^{-8} s, this will generate higher values of T_2. Narrowing of the corresponding resonances, typ-

* This work was supported by funds from the Institut Pasteur and from the CNRS UA 1129 and 1444. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed. Tel.: 331-45688872; Fax: 331-45688885; E-mail: alecrois@pasteur.fr.

[1] Martineau, P., Leclerc, C., and Hofnung, M., (1997) *Mol. Immunol.,* in press.

This paper is available on line at http://www-jbc.stanford.edu/jbc/

This is an Open Access article under the CC BY license.

ically to 5–20 Hz, will lead to the emergence of the corresponding cross-peaks from the background and to their specific observation. NMR dynamic filtering can thus be used to detect selectively regions of greater mobility in large proteins.

NMR studies were therefore performed to investigate the flexible regions of the MalE protein, of the insertion/deletion mutants, and of the C3 hybrids in order to determine if the insertion and deletion modifications of the structure of the wild-type protein affected its mobility. Flexibility of the C3 epitope in the different hybrid proteins was compared and analyzed with respect to the insertion site localization. This study is part of a more general work on the structural and immunological characteristics of heterologous peptides inserted into the MalE vector and on a potential correlation between these properties and the peptide environment in the hybrid proteins.[1,2,3]

EXPERIMENTAL PROCEDURES

Expression of Mutant Proteins—Insertion/deletion mutants corresponding to *Bam*HI sites introduced after residues 133, 161, 206, 211, 285, 296, 303, 339, and 363 were obtained as described previously (5). Hybrid proteins were prepared by insertion of an oligonucleotide coding for the C3 peptide into the previously introduced *Bam*HI sites (11). Except for site 339, all these insertion sites were permissive.[1] The C3 peptide was also inserted into the natural *Bcl*I nonpermissive site located after residue 348.[1]

Mutant genes were expressed on multicopy plasmids derived from pBR322 in *Escherichia coli* strain PD28, which carries a nonpolar deletion in the chromosomal *malE* gene (Δ*malE*444) and is constitutive for *malE* gene expression (*mal*T[c1]) (4, 15). Cultures were grown in LB medium (16) supplemented with 0.2% maltose and 100 μg/ml ampicillin for 7 h at 30 °C. Final A_{600} was between 2 and 3 ($1–2 \times 10^9$ bacteria/ml). MalE proteins were released from the periplasm by osmotic shock (17).

Purification—All proteins, except the insertion/deletion mutant MalE339, the corresponding hybrid MalE339C3, and the hybrid MalE348C3, in which C3 is introduced at a natural *Bcl*I site, were purified by affinity chromatography on a cross-linked amylose column according to Ferenci and Klotz (18). The MalE339, MalE339C3, and MalE348C3 proteins, which were unable to bind to the amylose column, were purified by ion-exchange chromatography on a MonoQ fast protein liquid chromatography column (Pharmacia Biotech Inc.) equilibrated in 20 mM Tris-HCl buffer, pH 7.5. Elution was performed with a linear gradient from 0 to 150 mM NaCl.[1] Purified MalE protein derivatives were precipitated with ammonium sulfate at 80% saturation and stored at 4 °C.

NMR Experiments—Proteins (~10 mg) were recovered by centrifugation; suspended in a minimum volume of 10 mM sodium phosphate buffer, pH 6.5; exhaustively dialyzed against the same buffer; and freeze-dried. Samples were dissolved in 650 μl of either H_2O/D_2O (9:1, v/v) or 99.96% D_2O (Euriso-top). Unless otherwise indicated, maltose was added to 1 mM. The final concentration of MalE proteins was 0.2–0.6 mM, depending on the sample, and that of the salt concentration was ~10 mM. When needed, acidification to pH 4.2 was performed by addition of diluted H_3PO_4.

Concentration of the wild-type MalE protein was calculated from its UV absorbance at 280 nm using an extinction coefficient of 68,750 M^{-1} cm^{-1}.[4] Mutants were quantified by colorimetric assay using the Bio-Rad kit and wild-type MalE as a standard (19). Homogeneity of the MalE proteins was checked on SDS-polyacrylamide gel (12%) before and after the NMR experiments (20).

NMR spectra were recorded at 35 °C on a Varian Unity 500 spectrometer equipped with Sun Sparc computers. TOCSY[5] experiments using MLEV17 to produce isotropic mixing (21, 22) were carried out with mixing times ranging from 20 to 80 ms. NOESY experiments (23) were recorded with mixing times of 60 and 100 ms. Saturation of the solvent proton resonance was accomplished by selective irradiation during the relaxation delay (2 s) and also during the mixing time in the NOESY experiments. Data sets were collected in the phase-sensitive mode with the hypercomplex scheme (24). Spectra were recorded with a sweep width of 7000 Hz in both dimensions and 32 scans/T_1 increment. A total of 512 free induction decays of 2048 data points were collected in T_2. Prior to Fourier transformation, spectra were zero-filled to final matrices of 2048×4096 points in T_1 and T_2, respectively, and matrices were multiplied by a shifted squared sine bell window function to enhance the resolution.

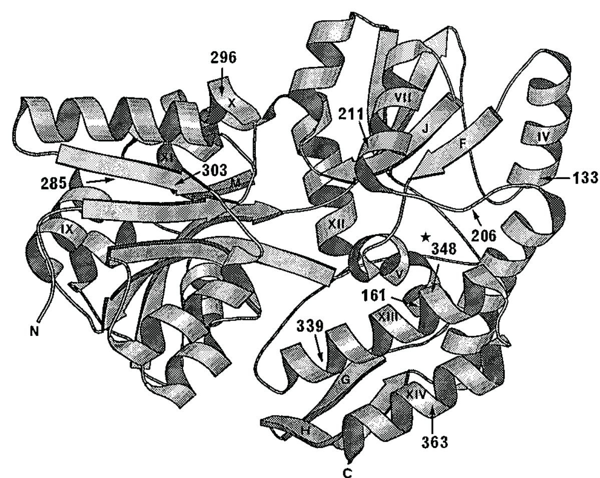

FIG. 1. **Location of C3 insertion sites on the structure of MalE.** The ribbon-style drawing was generated using Molscript (26). Secondary structure elements were determined according to Kabsch and Sander (27). *Arrows* indicate approximately the beginning of the insertion/deletion mutants. The position labeled with an *asterisk* corresponds to the *Bcl*I natural restriction site of the *malE* gene. Numbering is that of Spurlino *et al.* (25).

RESULTS

The overall structure of MalE consists of two distinct globular domains separated by a deep groove, in which maltose binds, and connected by three short segments. Both domains exhibit similar supersecondary structure: a core of β-sheets flanked on both sides with α-helices (Fig. 1) (25). Secondary structure element characterization according to Kabsch and Sander (27) gives 50% of the residues in α-helices, 22% in β-sheets, and the remaining 28% in loops and coils. As expected for a molecule of this size, no flexible region was detected by NMR with our criteria in the wild-type protein. TOCSY and NOESY spectra of the protein in H_2O/D_2O showed only a few cross-peaks. Signal intensity was very low, and these signals could hardly be distinguished from the background, especially in the NH-αH region. The lack of cross-peaks in this region was not due to an exchange of the amide protons with deuterium since spectra recorded at pH 6.5 and 4.2 were similar. Only a few spin systems could be determined, and corresponding residues did not fit with any continuous sequence in the protein.

As with the wild-type protein, MalE mutants showed spectra with very low amide proton signals in H_2O/D_2O (9:1, v/v). Therefore, only TOCSY experiments in D_2O were performed to investigate the MalE protein flexible regions. Fingerprints of the insertion/deletion mutants, of the corresponding C3 hybrids, and of wild-type MalE were compared three by three and carefully analyzed. The cross-peaks with significant intensity that were not present in the wild-type MalE spectra and that appeared in the mutant spectra were identified to known spin systems. Reconstitution of sequences present in the inserts and in the vector was attempted from the corresponding residues. Such sequences determining mobile segments were found in all the hybrids and in some of the insertion/deletion proteins. To discard the possibility that some signals of the protein and of the maltose overlap, preventing the identification of other residues, spectra of samples without maltose were recorded. No

[2] Saul, F. A., Vulliez-le-Normand, B., Lema, F., and Bentley, G. (1997) *Proteins Struct. Funct. Genet.*, in press.

[3] Vulliez-le-Normand, B., Saul, F. A., Martineau, P., Hofnung, M., Lema, F., and Bentley, G. (1997) *Protein Eng.,* in press.

[4] J. M. Betton, personal communication.

[5] The abbreviations used are: TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser effect correlation spectroscopy.

TABLE I

Location of C3 insertion sites on the primary and secondary structures of MalE and corresponding sequences in the insertion/deletion mutants and the C3 hybrid proteins

Secondary structure elements were determined according to Kabsch and Sander (27).

```
                                  αIV                       βF
MalE       P K T W E E I P A L D K E L K A K G K S A L M F
MalE 133   P K T W E E I P             D P     G K S A L M F

                αV                βG              βH
MalE       W P L I A A D G G Y A F K Y E N G K Y D I K D V
MalE 161   W P L I                 R I R               I K D V

                                  αVII                  βJ
MalE       K H M N A D T D Y S I A E A A F N K G E T A M T
MalE 206   K H M N A         R I R         F N K G E T A M T
MalE 211   K H M N A D T D Y S         G S         E T A M T

              αIX                        αX
MalE       L A K E F L E N Y L L T D E G L E A V N K D K P
MalE 285   L A K E F L E N YLLRIRTD E G L E A V N K D K P

             αX               βM             αXI          αXII
MalE       A V N K D K P L G A V A L K S Y E E E L A K D P
MalE 296   A V N K D               T D P               K D P
MalE 303   A V N K D K P L G A V A     P D P     E L A K D P

                        αXIII                          αXIV
MalE       S A F W Y A V R T A V I N A A S G R N T V D E A
MalE 339   S A F                     S D P                 D E A
MalE 348   S A F W Y A V R T A V I/N A A S G R N T V D E A
                               D N P A S T T N K D K G I

                       αXIV
MalE       D E A L K D A Q T R I T K
MalE 363   D E A L K D               G S E C C E M P D A A
```

TABLE II

Flexible sequences in the insertion/deletion mutants and in the C3 hybrid proteins

Residues observed by NMR are indicated in boldface letters; residues not observed are in lightface letters. Residues in italic correspond to oneface or more residues of the same type. Underlined lightface residues are uncertain.

Protein	N-flank	Insert	C-flank
MalE	EEIP/	ALDKELKAK	/GKSAL
133C3	EEIP/	DP/**DNPASTTN*KDK*/D**P	/GKSAL
133	EEIP/		/GKSAL
MalE	WPLI	AADGGYAFKYENGKYD	/IKDVK
161C3	WPLI/	RI/DNP**ASTTN*KDK*/**G**IR**	/IKDVK
161	WPLI/	RIR	/IKDVK
MalE	HMNA/	DTDYSIAEAA	/FNKGE
206C3	HMNA/	RI/**DNPASTTN*KDK***/GIR	/FNKGE
206	HMNA/	RIR	/FNKGE
MalE	TDYS/	IAEAA	/ETAMT
211C3	TDYS/	G**S/DNPASTTN*KDK*/**GS	/ETAMT
211	TDYS/	GS	/ETAMT
MalE	NYLL/		/TDEGL
285C3	NYLL/	RI/**DNPASTTN*KDK***/G**IR**	/**T**DEGL
285	NYLL/	RIR	/TDEGL
MalE	VNDK/	KPLGAVALKSYEEELA	/KDPRI
296C3	VNDK/	TDP/**DNPASTTN*KDK***/DP	/KDPRI
296	VNDK/	TDP	/KDPRI
MalE	GAVA/	LKSYEE	/ELAKD
303C3	GAVA/	PDP/**DNPASTTN*KDK*/DP**	/**ELA**KD
303	GAVA/	PDP	/ELAKD
MalE	MSAF/	WYAVRTAVINAASGRTV	/DEALK
339C3	M**SAF/**	**SDP/DNPASTTN*KDK*/DP**	**/DEALKDAQTRIT*K***
339	M**SAF/**	***SDP***	**/DEALKDAQTRIT*K***
MalE	TAVI/		/NAA
348C3	TAVI/	DNP**ASTTN*KDK*/*GI***	/**NAAS**G**RQTVDEAL*K*DAQT*R*IT*K***
MalE	ALKD/	AQTRITK	
363C3	ALKD/	G**S/DNPASTTNKDK**/G**S**	/***E***CC**EMPDAA**
363	ALKD/	G**S**	/**E**CC**EMPD***AA*

significant difference with spectra of samples containing 1 mM maltose was observed. The location of the insertion sites within the secondary structure of MalE is shown in Table I, and results on the respective mobility of the MalE proteins are presented Table II.

Insertion Site 133—The insertion/deletion mutant MalE133 comprises a deletion of 9 residues in helix IV at the surface of the C domain and an insertion of 2 residues, DP. The corresponding hybrid MalE133C3 presents an insertion of 15 residues, the 11 residues of the C3 insert and the 4 residues of the flanking sequences, DP and DP. The MalE133 spectrum did not show any significant change as compared with the MalE spectrum, whereas in MalE133C3, new cross-peaks appeared (Fig. 2). One proline, 1 alanine, 1 serine, and 2 threonines could be unambiguously identified from their characteristic spin systems and their chemical shifts (28). Lysine and five spin systems which could agree with aspartic acid or asparagine were also detected. The only sequence that could be reconstituted from these residues in the hybrid protein was therefore DNPNKDKD, provided both lysines had approximately similar chemical shifts. This sequence corresponds to that of the C3 epitope, DNPNKDK, extended with the first residue of its C terminus flanking sequence, DP.

Although weakly packed on the central β-sheet in the C domain, helix IV was not found to be mobile in MalE. Deletion of three-quarters of its residues did not reveal any flexible region in the insertion/deletion mutant MalE133. It has already been suggested that α-helix deletion accommodation in insertion/deletion mutants occurred by shifting the packing and bonding of both the small inserted linker and the vector residues on each side of the insert (6). In view of the pitch per residue between an α-helix (1.5 Å) and a β-strand (3.5 Å), one proposed mechanism to compensate for a deletion in a helix was the creation of a strand. Our results are compatible with this hypothesis. On the contrary, when replaced by peptide C3 in the MalE133C3 hybrid, helix IV changes for a flexible segment of 12 residues, the 11 residues of C3 and 1 residue of the flanking sequence. Peptide C3 has no intrinsic propensity to form a helix and corresponds to a loop in the original poliovirus VP1 protein. Due to its number of residues, it would not be able to form a strand without major changes in the vector structure when inserted into MalE and should therefore form a loop. The loop, too long to interact with the C domain central β-sheet, would move freely. Results show that the flanking regions of the vector protein, a turn of helix on one side and a small coiled segment on the other, are not destabilized.

Insertion Site 161—In MalE161, which lacks the entire G strand and part of the H strand of the C domain two-stranded antiparallel GH β-sheet plus a part of helix V, 16 residues are

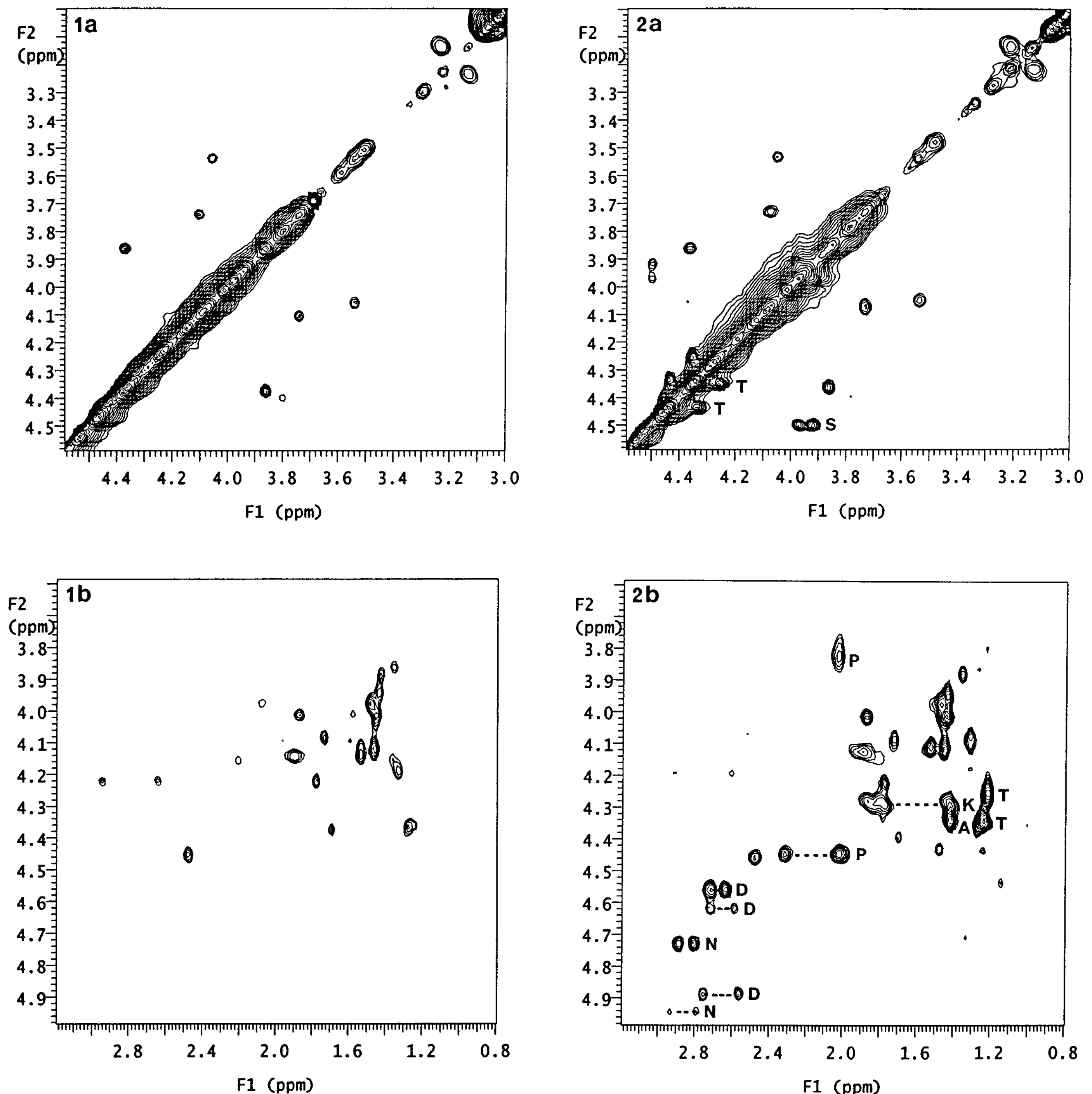

FIG. 2. **Regions of the MalE133 (*panels 1*) and MalE133C3 (*panels 2*) TOCSY spectra presenting significant differences.** *Panels a*, αH-αH; *panels b*, αH-βH, αH-γH, and αH-δH regions. Experiments were performed in D_2O at 35 °C and pH 6.5 with a mixing time of 60 ms. See "Discussion" for the assignment of Asn and Asp.

deleted and 3 inserted (GIR). The corresponding MalE161C3 hybrid comprises a 16-residue insertion, including the flanking sequences RI and GIR.

Spectra of MalE161 and of the wild-type protein were rather similar. Differences principally occurred in the aromatic region, where some cross-peaks disappeared in the mutant. As the deleted sequence contains 3 tyrosines (positions 167, 171, and 176) and 1 phenylalanine (position 169), missing signals are probably related to some of these residues. In spite of the length of the deletion, no flexible region became apparent. This suggests an important shift of the packing and bonding of the linker and of the vector flanking residues together with the existence of interactions between these residues and both helices XIII and XIV.

The sequence ASTTNKDKGIR could be unambiguously reconstituted from the identified residues in the spectrum of MalE161C3. The fact that no glycine was observed indicates that its two α-protons have similar chemical shifts, which is in agreement with a great flexibility of this residue. This result shows that the first 3 residues of C3 (DNP) present tight interactions with the other structural elements of the C domain. The fact that helix V extends to the linker RI and to these residues may not be excluded. As in the MalE133C3 hybrid, other deleted secondary structure elements are replaced by a flexible loop, probably without consequence on the tertiary structure of the vector.

Insertion Sites 206 and 211—MalE206 and MalE211 comprise deletions of 10 or 9 residues, respectively, in helix VII, which faces the binding cleft in the C1 domain, and insertions of 3 (RIR) or 2 (GS) residues, respectively. Spectra of the two

mutants did not reveal any mobile region. Although helix VII is well packed on the central β-sheet of the C domain, removal of its first two (MalE206) or last two (MalE211) turns did not induced any flexible region in the molecule. It is likely that accommodation of the deletion within helix VII occurs with the same mechanism as that suggested for the MalE133 mutant.

Spectra of both the MalE206C3 and MalE211C3 hybrids allowed the reconstitution of the C3 epitope sequence, with 2 extra serine residues most probably belonging to sequences flanking the insert in MalE211C3. It appears from these results that C3 peptides are inserted in positions 206 and 211 as loops flexible throughout their length and that their insertion does not produce the unfolding of the remaining parts of helix VII or their unpacking from the C domain central β-sheet. A similar observation was made for helix IV in the MalE133C3 hybrid.

Insertion Site 285—MalE285 presents no deletion, but only an insertion of 3 residues (RIR) in the coiled segment connecting the two helices IX and X of the N domain. No mobile region was observed in this mutant.

In the corresponding MalE285C3 hybrid, 13 residues fitting in the sequence DNPNKDKGIRT were unambiguously detected. Since a few residues of the vector protein on one side at least of insertion site 285 had to move to allow the C3 insertion, it is therefore likely that the opening of the main chain is just large enough to tolerate the insertion. Indeed, 1 residue only of the vector (Thr^{286}) became mobile, and helices IX and X did not unfold.

Insertion Sites 296 and 303—Deletions in MalE296 and MalE303 are located in the same region of the N domain, but they widely differ in length. MalE296 comprises deletions of β-strand M and of helix XI, *i.e.* 16 residues, and includes a TDP insertion. MalE303 lacks only a part of helix XI, 6 residues, and includes a PDP insertion. Spectra of the insertion/deletion mutants did not show any mobile region.

In the MalE296C3 hybrid, all the C3 residues were identified, but none of the flanking sequences, TDP and DP. In the MalE303C3 hybrid, the sequence DNPASTTNKDKDPELA could be reconstituted from the residues identified, *i.e.* the C3 epitope, its C-terminal flanking sequence, and the first 3 residues of the vector on the same side. An ambiguity remained about the mobility of the 4th residue of the vector Lys^{313} because of the overlapping of the lysine cross-peaks. Insertion of epitope C3 as a loop in site 296 did not destabilize the vector protein, although two of its secondary structure elements, β-strand M and helix XI, were deleted. In MalE303C3, however, in which only a part of helix XI was suppressed, insertion of peptide C3 made the remaining last turn of helix flexible, most probably via its unfolding. Such a destabilization had not been observed in the MalE133C3, MalE206C3, and MalE211C3 hybrids, although deletions also involved parts of helices.

Insertion Site 339 (Nonpermissive)— MalE339 comprises a deletion of 17 residues, including the last three turns of helix XIII and the segment connecting helix XIII to the C-terminal helix XIV, and an insertion of 3 residues, SDP. The spectrum showed that its C-terminal region was entirely flexible, from Ser^{337} to Lys^{370}, which indicates that the remaining first turns of helix XIII and of helix XIV were destabilized and have most probably lost their secondary structure.

As with MalE339, the corresponding MalE339C3 hybrid was flexible from residue 337 to the C terminus of the molecule. The C3 epitope, both flanking sequences SDP and DP, and 15 residues of the vector, 3 on the N-terminal side of the insert and 11 on the C-terminal side, were identified, indicating the unfolding of both helices XIII and XIV as in the insertion/deletion mutant. Cross-peaks related to the C-terminal 358–370 residues were superimposable in both MalE339 and MalE339C3.

Insertion Site 348 (Nonpermissive)—As MalE348C3 was obtained by insertion of the C3 coding oligonucleotide in a natural restriction site of the *malE* gene, there was no corresponding insertion/deletion mutant, and the spectrum of the hybrid was directly compared with that of the wild-type protein. MalE348C3 comprises an insertion of 13 residues in the last turn of helix XIII, the C3 epitope extended with a GI flanking sequence on its C-terminal side. The spectrum of the hybrid protein presented a large number of cross-peaks, which allowed the reconstitution of a sequence including the last 7 residues of the epitope (STTNKDK), the flanking sequence GI, and the entire C-terminal region of the vector NAASGRQTVDEALKDAQTRITK. An ambiguity remained about the C3 alanine residue since 4 more alanines were present in the mobile region and the corresponding cross-peaks overlapped. Chemical shifts of residues 358–370 were close to those of the same residues in the MalE339 and MalE339C3 mutants. In the MalE348C3 hybrid, the first 3 or 4 N-terminal residues of C3, DNP(A), are not flexible. On the other hand, the vector completely unfolds on the C-terminal side of the insert. We can imagine that the last turn of helix XIII corresponding to the sequence NAAS is withdrawn from the helix by the insertion of C3, but is replaced by another turn formed from the DNP or DNPA sequence of the insert. As with the sequence NAAS, the sequence DNP(A) would therefore be quite motionless. Insertion accommodation of the other residues of C3 and of the NAAS sequence between helices XIII and XIV would induce the destabilization of helix XIV and its entire unfolding.

Insertion Site 363—Deletion in the insertion/deletion mutant MalE363 occurs in the C-terminal region of the MalE protein and comprises the last turns of helix XIV. Insertion involves 2 residues, GS, encoded by the *Bam*HI linker and 9 residues, ECCEMPDAA, encoded by an open reading frame of the intergenic *malE-malF* region due to the deletion of the stop codon of the *malE* gene. The inserted sequence could be reconstituted from residues identified in the spectra. The 2 sequential cysteines, not observable, are likely to be implicated in a disulfide bridge, which makes them less flexible. In MalE363C3, the segment including the flanking sequence GS, epitope C3, and the open reading frame of the intergenic region appeared to be mobile. Chemical shifts of residues implicated in that region were similar in both the insertion/deletion and hybrid proteins. The fact that no residue of the vector C-terminal region could be detected in MalE363 and in MalE363C3 indicates that the two remaining turns of helix XIV were not destabilized in either of the mutants.

Both MalE363 and MalE363C3 proteins are able to bind to an amylose column, in contrast to MalE339, MalE339C3, and MalE348C3. The MalE339 and MalE339C3 proteins lack Trp^{340}, an amino acid crucial for the binding of maltose (29). An unfolding of their C-terminal helices XIII and XIV was shown by NMR. In MalE348C3, while helix XIII remains structured, helix XIV unfolds completely. In the MalE363 and MalE363C3 proteins, which present a deletion of the two C-terminal turns of helix XIV, helix XIII and the three first turns of helix XIV remain stable. This observation is in agreement with previous data showing that deletion of both helices XIII and XIV resulted in a loss of the maltose binding activity (5).

DISCUSSION

NMR studies of the different MalE hybrid proteins showed that, wherever its insertion site in the MalE vector, peptide DNPASTTNKDK was fully or mostly flexible: 3 residues (DNP) were not observed in hybrids MalE161C3 and MalE348C3, and in all the other C3 mutants, the 11 residues were detected. This

flexibility is not restricted only to the side chains, but concerns the whole inserted fragment as spin systems could be reconstituted from αH protons up to the end of the chain. Based on line width analysis, the C3 epitope motions are therefore at least 10 times faster than the overall tumbling of the molecule. This result is in agreement with x-ray data on MalE133C3. Whereas structures of the MalE part of the hybrid protein and of MalE were essentially similar, the C3 epitope was too mobile to be seen in the crystal (30).

The C3 insertion sites are located within various elements of secondary structure of the MalE protein, the α-helix, β-sheet, and coiled segment. Although on the surface, these elements are tightly packed with the rest of the protein since no flexible regions were noticed in the wild-type protein. Their partial or total deletion did not produced any unpacking of the molecule, except in MalE339, in which deletion involves parts of both C-terminal helices XIII and XIV. Insertion of peptide C3 in place of the deleted residues or merely inside a coiled segment (MalE285C3) or a helix (MalE348C3) destabilized the vector protein only when the insertion site was located in helix XIII (MalE339C3 and MalE348C3). The inserted residues thus adopt a loop conformation in hybrid proteins MalE133C3 to MalE303C3. In hybrid proteins MalE339C3, MalE348C3, and MalE363C3, they are included in the flexible C-terminal part of the molecule. As shown by Martineau *et al.*,[1] all the hybrid proteins bind to an anti-C3 monoclonal antibody with a highly increased antigenicity when compared with the free peptide, except MalE339C3, which does not present any detectable activity for the monoclonal antibody. In both MalE339C3 and MalE348C3, the vector protein unfolds on the C-terminal side of the insert, and in MalE363C3, the open reading frame of the intergenic region is mobile. However, in MalE348C3 and in MalE363C3, the C3 insert is constrained on its N-terminal side, whereas in MalE339C3, it is totally unrestrained and likely able to adopt a range of conformations similar to that of the free C3 peptide. This excess of mobility would explain the lack of antigenicity of the MalE339C3 hybrid.

The flanking sequences on the C-terminal side of the insert were observed in most of the hybrids, without an obvious correlation between their composition and their intrinsic conformational freedom. The sequence GS, a most hydrophilic and flexible sequence, was indeed observed in both MalE211C3 and MalE363C3, but the sequence GIR, which was not seen in the MalE206C3 hybrid, was identified in MalE161C3 and MalE285C3. And while the sequence DP, which contains the most constrained amino acid, proline, was not detected in MalE296C3, it was partly observed in MalE133C3 and totally in MalE303C3 and MalE339C3. It should be noted, however, that in the MalE339C3 hybrid, DP is included in a flexible region. The flanking sequence on the N-terminal side of the inserted peptide was only detected in two or three proteins: hybrid MalE363C3 and possibly hybrid MalE211C3, in which this sequence was GS, and hybrid MalE339C3, for which insertion was shown to destabilize the vector protein on both sides of the insert.

Residues Asn^{94}, Pro^{95}, Ser^{97}, and Asn^{100} of the C3 insert were assigned in all the MalE hybrids. These residues were represented only once in some of the observed sequences, as in MalE161C3, and their proton chemical shifts were so constant from one hybrid to another that they could be unambiguously identified in the different proteins. Ala^{96} was attributed to hybrids MalE133C3 to MalE303C3 only because of the overlap of its αH-βH cross-peak with those of the other alanines also observed in hybrids MalE339C3 to MalE363C3. Assignment of both threonines was arbitrarily performed, with position 98 being given to the one presenting the lower αH chemical shift.

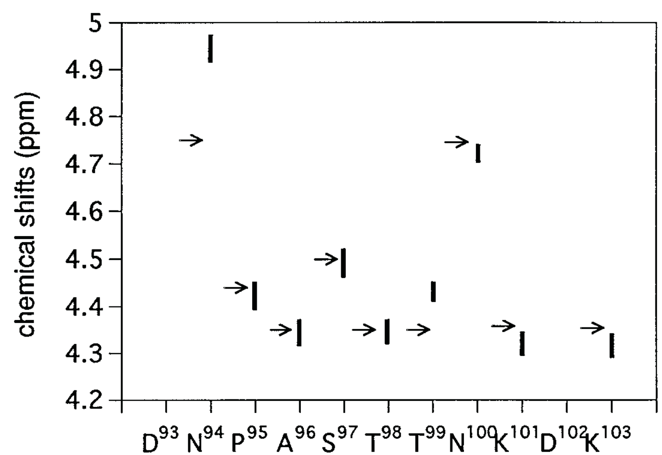

FIG. 3. **Chemical shifts of the C3 residue αH protons in the hybrid proteins.** All insertion sites were considered for Asn^{94}, Pro^{95}, Ser^{97}, Asn^{100}, Lys^{101}, and Lys^{103}. Sites 133–296 are for Ala^{96} and Thr^{99}. Sites 133–363 are for Thr^{98}. Numbering of the C3 residues is that of Hogle *et al.* (13). *Arrows* indicate the chemical shifts expected for random coil residues (28).

With this agreement, Thr^{98} was assigned in all the C3 hybrids and Thr^{99} in hybrids MalE133C3 to MalE303C3 only, with the other threonines observed in these proteins presenting too similar chemical shifts. Lys^{101} and Lys^{103}, which always merged into one another, were identified in all the hybrids. Asp^{93} and Asp^{102} could not be assigned since the αH and βH chemical shifts of most of the aspartic acids observed in the spectra of the different hybrids were very close, respectively. As shown Fig. 3, the αH chemical shifts of the C3 residues are close to those expected for random coil residues (28). The 4.92–4.97 ppm values observed for Asn^{94} indeed correspond to the fact that αH chemical shifts of residues located before a proline are significantly shifted downfield (31). Aspartic acids of the flexible DP flanking sequences of hybrids MalE133C3, MalE303C3, and MalE339C3 could be assigned for the same reason since, due to their positions, their αH chemical shift values (4.83–4.89 ppm) highly exceeded the values determined for the other observed aspartic acids (4.54–4.62 ppm). The C3 residue chemical shift data are thus in agreement with the flexible loop or coiled segment conformation suggested above. Differences observed between Thr^{98} and Thr^{99} α- and β-protons are more likely to be due to a sequential than to a structural effect since they were observed when the peptide was tightly maintained at both its extremities as well as when it was freely moving with the C-terminal region of the MalE vector.

The crystal structure of the complex between the peptide corresponding to residues 86–103 of the VP1 protein and the Fab fragment of a neutralizing monoclonal antibody raised against the heat-inactivated virion was recently determined at 3-Å resolution (32). Residues 93–103 corresponding to peptide C3 adopt mostly an S-shaped conformation, with 3 residues in an extended conformation (residues 93–95) and two type I β-turns in series (residues 96–103). Only residues 96–103 directly contact the Fab fragment. Residues 96–99 (ASTT) adopt an identical conformation in the peptide and in the corresponding region of the VP1 protein, with main chains superimposing with a root mean square of 0.68 Å. The conformation of residues 93–95 is somewhat similar, but that of residues 100–103 is substantially different. When the Fab fragment was docked on the virus VP1 protein by superimposition of the 4 residues ASTT in the complex and in the VP1 protein, many unfavorable interactions occurred between the Fab fragment and the VP1 C3 loop C-terminal region (from C3 residue Asn^{100}). Furthermore, the docked model was not completely consistent with the

virus binding and neutralization. Wien *et al.* (32) therefore proposed a two-step binding process in which the antibody induces conformational changes in the C3 loop after attachment to the virus, which would allow the C-terminal part of the loop to adopt the conformation seen in the complex. When inserted in MalE at positions 133 and 303, the C3 peptide is antigenic (11). It is detected by a polyclonal antibody as well as by a monoclonal antibody recognizing the C3 epitope in the native virion. Insertion in most of the other sites also produces antigenic hybrids.[1] It is likely that in these hybrids, the C3 peptide adopts a conformation similar or at least close to that present in the complex. In contrast to residues 93–95, residues 96–103, forming the real epitope, as determined in the peptide-Fab complex structure, were found to be flexible in all the hybrids we studied. This flexibility, often increased in the peptide C-terminal region by that of the C-terminal flanking sequence, would allow the C3 epitope to adapt its conformation to that of the antibody.

REFERENCES

1. Arnon, R., and Sela, M. E. (1969) *Proc. Natl. Acad Sci. U. S. A.* **62,** 163–170
2. Vuillemier, S., and Mutter, M. (1992) in *Structure of Antigens* (Van Regenmortel, M. H. V., ed) pp. 79–128, CRC Press, Inc., Boca Raton, FL
3. Hofnung, M., Bedouelle, H., Boulain, J.-C., Clement, J.-M. C., Charbit, A., Duplay, P., Gehring, K., Martineau, P., Saurin, W., and Szmelcman, S. (1988) *Bull. Inst. Pasteur* **86,** 95–101
4. Duplay, P., Bedouelle, H., Fowler, A., Zabin, I., Saurin, W., and Hofnung M. (1984) *J. Biol. Chem.* **259,** 10606–10613
5. Duplay, P., Szmelcman, S., Bedouelle, H., and Hofnung, M. (1987) *J. Mol. Biol.* **194,** 663–673
6. Betton, J. M., Martineau, P., Saurin, W., and Hofnung, M. (1993) *FEBS Lett.* **325,** 34–38
7. Stahl, S., Sjolander, A., Nygren, P.-A., Berzins, K., Perlmann, P., and Ulhen, M. (1989) *J. Immunol. Methods* **124,** 43–52
8. Leclerc, C., Martineau, P., van der Werf, S., Deriaud, E., Duplay, P., and Hofnung, M. (1990) *J. Immunol.* **144,** 3174–3182
9. O'Callaghan, D., Charbit, A., Martineau, P., Leclerc, C., van der Werf, S., Nauciel, C., and Hofnung, M. (1990) *Res. Microbiol.* **141,** 963–969
10. Martineau, P., Charbit, A., Leclerc, C., Werts, C., O'Callaghan, D., and Hofnung, M. (1991) *Bio/Technology* **9,** 170–172
11. Martineau, P., Guillet, J.-G., Leclerc, C., and Hofnung, M. (1992) *Gene* (*Amst.*) **113,** 35–45
12. Wychowski, C., van der Werf, S., Siffert, O., Crainic, R., Bruneau, P., and Girard, M. (1983) *EMBO J.* **2,** 2019–2024
13. Hogle, J. M., Chow, M., and Filman, J. D. (1985) *Science* **229,** 1358–1365
14. Jardetzky, O., and Roberts, G. C. K. (1981) *NMR in Molecular Biology*, pp. 448–492, Academic Press, New York
15. Shuman, H. A. (1982) *J. Biol. Chem.* **257,** 5455–5461
16. Miller, J. H. (1972) *Experiments in Molecular Genetics*, p. 11724, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
17. Bedouelle, H., and Duplay, P. (1988) *Eur. J. Biochem.* **171,** 541–549
18. Ferenci, T., and Klotz, U. (1978) *FEBS Lett.* **94,** 213–217
19. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
20. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G., and Struhl, K. (1993) *Current Protocols in Molecular Biology*, p. 451, John Wiley & Sons, Inc., New York
21. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson.* **65,** 355–360
22. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. R. (1988) *J. Am. Chem. Soc.* **110,** 7870–7872
23. Kumar, A., Ernst, R. R., and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun.* **95,** 1–6
24. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magn. Reson.* **48,** 286–292
25. Spurlino, J. C., Lu, G.-Y., and Quiocho, F. A. (1991) *J. Biol. Chem.* **266,** 5202–5219
26. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
27. Kabsch, W., and Sander, C. (1983) *Biopolymers* **22,** 2577–2637
28. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, p. 17, John Wiley & Sons, Inc., New York
29. Martineau, P., Szmelcman, S., Spurlino, J. C., Quiocho, F. A., and Hofnung, M. (1990) *J. Mol. Biol.* **214,** 337–352
30. Rodseth, L. (1992) *Crystallographic Studies of the Maltodextrin-binding Protein from Escherichia coli.* Ph.D. thesis, Rice University
31. Wishart, D. S., Bigam, C. G., Holm, A., Hodjes, R. S., and Sykes, B. D. (1995) *J. Biomol. NMR* **5,** 67–81
32. Wien, M. W., Filman, D. J., Stura, E. A., Guillot, S., Delpeyroux, F., Crainic, R., and Hogle, J. M. (1995) *Nat. Struct. Biol.* **2,** 232–243