The Crystal Structure of the Fab Fragment of the Monoclonal Antibody MAK33

IMPLICATIONS FOR FOLDING AND INTERACTION WITH THE CHAPERONE BiP

Received for publication, June 15, 2000, and in revised form, October 13, 2000
Published, JBC Papers in Press, October 17, 2000, DOI 10.1074/jbc.M005221200

John G. Augustine§§, Agustin de la Calle§§, Gerhard Knarr¶, Johannes Buchner††, and Christin A. Frederick∗∗

From the ∗Dana-Farber Cancer Institute and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115 and the †Institut für Organische Chemie und Biochemie, Technische Universität München, Lichtenbergstrasse 4, Garching 85747, Germany

The Fab fragment of the murine monoclonal antibody, MAK33, directed against human creatine kinase of the muscle-type, was crystallized and the three-dimensional structure was determined to 2.9 Å. The antigen-binding surface of MAK33 shows a convex overall shape typical for immunoglobulins binding large antigens. The structure allows us to analyze the environment of cis-prolyl-peptide bonds whose isomerization is of key importance in the folding process. These residues seem to be involved with not only domain stability but also seem to play a role in the association of heavy and light chains, reinforcing the importance of β-strand recognition in antibody assembly. The structure also allows the localization of segments of primary sequence postulated to represent binding sites for the ER-specific chaperone BiP within the context of the entire Fab fragment. These sequences are found primarily in β-strands that are necessary for interactions between the individual domains.

MAK33 is a murine monoclonal antibody of subclass κIgG1 specific for muscle-type creatine kinase (CK-MM), a key enzyme involved in the formation of creatine-phosphate in muscle cells. This antibody-antigen interaction is specific for the native dimeric form of the antigen only and results in the inhibition of the enzymatic activity, thus making MAK33 a valuable tool in the diagnosis of myocardial infarction (1). In recent years, the Fab fragment of MAK33, consisting of the entire light chain and the two N-terminal domains of the heavy chain, has been used as a model for studies of protein folding for multidomain proteins (2–5). Slow steps in the folding process of the MAK33 Fab were found to be due to isomerizations of prolyl-peptide bonds, however, the structural basis for these phenomena remained unknown. Folding catalysts of the peptidyl prolyl isomerase (PPIase) family have been shown to catalyze these slow folding processes in vitro (2). The resultant isomeric state in turn has been shown to influence subunit association. Kinetic analyses of the folding pathway have demonstrated that the relative position of polypeptide chain association within the folding pathway leading to the native molecule is different for the Fab fragment and the C3–C3 dimer of the Fc fragment (4, 6).

Although the isomerization of key proline residues could occur at various stages relative to subunit association for the Fab (4), dimer prolyl isomerization had to take place before association in the case of the C3–C3 (6). These results indicate a complex relationship between prolyl isomerization and intermolecular association in the folding of the antibody molecule.

During the biosynthesis of antibodies, the ER-specific hsp70 chaperone family member, BiP, plays an important role in the assembly of the mature antibody. BiP has been shown to bind to the variable domain of certain light chains to assist in folding or as a means of removing improperly folded chains from the secretory pathway (7). In addition, binding to the newly synthesized heavy chain allows the polypeptide chain to remain soluble until the appropriate light chain is translated and available to form the complete antibody molecule (8, 9).

Based on the screening of a phage display peptide library for BiP-binding sequences (10), an algorithm has been developed that allows for the prediction of potential protein sequences, which might constitute BiP-binding sites in proteins (11, 12). Analysis of the primary sequence of MAK33 as well as that of a related antibody indicated that binding does seem to require exposed hydrophobic residues. Notably leucine, but also tryptophan, residues are specifically enriched in BiP-binding sequences compared with the overall protein sequence. As evidenced from the stimulation of the ATPase activity of BiP upon heptapeptide binding in vitro, potential BiP-binding sites in vivo have been identified (11, 12).

In the present study, we have determined the three-dimensional structure of the MAK33 Fab fragment by x-ray crystallography. This structure enables us to discuss the protein environment of the cis-proline peptide bonds as well as the position of BiP-binding sites identified biochemically using synthetic heptapeptides. In this way one can begin to address the critical questions of subunit recognition within the context of the complete folding pathway of a multidomain protein.
**EXPERIMENTAL PROCEDURES**

**Protein Preparation and Crystallization**—The protocolically derived Fab fragment of the murine monoclonal antibody MAK33 of subclass κIgG1 directed against dimeric muscle-specific human creatine kinase (CK-MM, E.C. 2.7.3.2) was obtained from Roche Molecular Biochemicals GmbH. The lyophilized Fab was resuspended to ~6 mg/ml in 150 mM NaCl by gently rocking at 4 °C for 1 h. The solution was then dialyzed overnight against MilliQ water and subsequently concentrated to 30 mg/ml using Amicon microconcentrators. The final protein concentration was determined spectrophotometrically using the extinction coefficient of A280 = 80,000 M−1 cm−1 (13).

Crystals of MAK33 Fab were grown at 4 °C by vapor diffusion in 4-μl sitting drops made up of equal volumes of protein solution and a reservoir solution containing 50 mM citrate, pH 5.6, 16% (w/v) polyethylene glycol 8000, 1.0 M NaCl, and 16% glycerol. Large (0.25 × 0.20 × 0.80 mm3), single crystals were obtained after 8–10 weeks. The crystallographic asymmetric unit contains two Fab molecules and ~53% solvent with a Matthews coefficient of 2.64 Å3/dalton and a cell volume of 264,105.8 Å3 (14).

**Data Collection and Structure Determination**—Diffraction data were collected at 4 °C from CuKα radiation generated by a Rigaku RTP500 RC rotating anode generator using a Marresearch imaging plate. Reflection files were indexed using the program DENZO (15) and subsequently scaled by the program SCALEPACK (16). A summary of data collection statistics is shown in Table I. The structure of MAK33 Fab was solved by the molecular replacement method using the anti-influenza virus hemagglutinin Ha1 κIgG2A Fab fragment 17/9 (17) as a search model in the program AMoRe (18). The sequence identity between the search model and MAK33 Fab was 98 and 53% identical for the constant and variable domains of the heavy chain, respectively, and 87 and 89% identical for the constant and variable domains of the heavy chain, respectively. A rotational search using the entire Fab fragment was carried out, resulting in a pair of clear solutions with correlation coefficients significantly higher than those of the next highest solutions. After an initial rigid body refinement based on this model, a 2Fo − Fc difference map was calculated yielding strong electron density, and the molecular packing was checked showing no conflicts between molecules related by translational symmetry. The elbow regions connecting the variable and constant domains of both the light and heavy chains could be easily discerned in the electron density map, yielding an intact molecule for structure refinement.

Nonhomologous residues in β-strands were mutated to alanine, and the loops connecting the strands, including the six CDRs, were completely deleted to minimize model bias. Structure refinement was carried out using the program X-PLOR (19) utilizing noncrystallographic symmetry. In the initial cycles, the six CDRs were deleted from the model. The elbow regions connecting the variable and constant domains of both the light and heavy chains were easily discerned in the electron density map, forming the loop-connecting strands A and B of the heavy chain constant domain, were also not included in refinement and model building due to weak electron density. The final model geometry was analyzed with PROCHECK (21). Overall refinement statistics and final model parameters are given in Table I. Only 2 residues (0.6% of total) had phi-psi angle combinations that fall just outside of allowed regions of the Ramachandran plot. These are both located in loop regions of the protein exterior. The final coordinates have been submitted to the Protein Data Bank and have been assigned PDB ID 1FH5 and Research Collaboratory for Structural Bioinformatics ID RCSB011583.

**BiP ATPase Activity Assays**—The binding of MAK33 peptides by BiP was monitored by the measurement of BiP ATPase activity as described previously (11). BiP used in this assay was purified from bovine pancreas as described (22), with the addition of gel filtration chromatography as a final purification step. The MAK33 peptides were synthesized using Fmoc (N-(9-fluorenyl)methoxycarbonyl)-protected amino acids as described previously (23). The standard assay contained 40 mM HEPES, pH 7.0, 2 mM MgCl2, 500 μM unlabeled ATP, 10 μM of [γ-32P]ATP and ~4 μg of BiP in a total volume of 20 μl. Following different times of incubation at 37 °C, 3-μl aliquots were removed, and the amounts of ATP and ADP were determined by thin-layer chromatography and liquid scintillation.

### Table I

| Parameter                              | Value         |
|----------------------------------------|---------------|
| Space group                            | P1            |
| Unit cell dimensions (a,b,c) (Å)       | 58.91, 65.66, 73.46 |
| α, β, γ (°)                           | 109.14, 106.94, 95.91 |
| Resolution range (Å)                  | 25.0–2.9      |
| Unique reflections                     | 20,143        |
| -fold redundancy                       | 3.9           |
| Completeness (%)                       | 93.7 (87.1)   |
| Rmerge (%)                             | 6.1 (27.9)    |
| Rmerge (%)                             | 10.0 (2.5)    |
| No. of protein atoms                   | 3159/molecule |
| Resolution range (Å)                  | 8.0–2.9       |
| Reflections used                      | 18,312        |
| Completeness (%)                      | 89.4          |
| R factor (%)                           | 19.8 (28.1)   |
| Root mean-squared deviations (Å2)      | 24.3 (32.9)   |
| Root mean-squared deviations (Å2)      | 0.014         |
| Average B value (Å2): all atoms       | 37.16         |
| Average B value (Å2): mainchain atoms | 36.09         |
| Average B value (Å2): sidechain atoms | 38.32         |
| Residues within allowed regions of Ramachandran plot (%) | 99.4 |

### Results

**Overall Structure**—The three-dimensional structure of the MAK33 Fab fragment was determined using the molecular replacement method (see Table I). Two independent molecules were identified related by a pseudo 2-fold rotation. The final electron density maps were of high enough quality in most regions such that we could identify several previous DNA sequencing errors within the heavy chain (1). The MAK33 Fab fragment is comprised of four typical immunoglobulin domains, two each from the light and heavy chains (Fig. LA). Each domain is characterized by a β-barrel structure consisting of two β-sheets connected by an internal disulfide bridge. The orientation of the two interchain domain pairs with respect to each other is described by the elbow angle between the pseudo- dyads V1-V3 and C1-C3-1 in MAK33. The two molecules within the asymmetric unit have very similar domain orientations, assuming typical elbow angles of 161.9° and 161.3° for molecules 1 and 2, respectively.

Surface loops extending from the N-terminal variable domains of both heavy and light chains form the antigen-binding site. The conformations of these loops can be classified according to the observations of Chothia and Lesk (24). They are typical for κIgG family members and belong to the same classes as those of the closely related antibody, 3D6, whose structure has been determined (25). The only possible exception is the H3 loop whose conformation appears disordered in our structure. Overall the surface comprised of the MAK33 CDR loops contains mostly neutral or uncharged residues (Fig. 1B). However, two strongly charged zones of only a few residues each are localized in the heavy chain H1 and H2 loops. The only charged surface potential found in the light chain is part of the CDR L2, whose surface is partially directed away from the main antigen binding area of the other CDRs (Fig. 1B).
This antibody is distinct in that it recognizes only the native dimeric form of its antigen, CK-MM, and in contrast to many other antibodies it does not bind the denatured form of this protein. Therefore, antigen recognition is expected to be a conformation-specific, as opposed to a sequence-specific, interaction, i.e. requiring either a complex epitope comprising residues from both subunits or one whose key residues are influenced by the dimer formation. Comparison of the antigen-binding surface of MAK33 with those described in the complex structures of two sequence-specific antibodies directed against hen egg lysozyme reveals certain subtle differences. The contact surface of HyHel-10, an α/IgG1 antibody, is a broad and relatively flat area with no obvious grooves or cavities but instead one extended region, comprising residues from CDRs H1 and H2, that appears to interact with the lysozyme active site (26). The CDR loop H3 is very short in this antibody and contributes only 1 residue to the interaction. The binding surface of monoclonal antibody F9.13.7, as seen in complex with a cross reacting antigen, Guinea fowl lysozyme, includes a shallow surface groove one side of which is mainly composed of the CDR H3 loop. In this antibody CDR H3 is relatively long and positioned such that its length and composition directly influence the binding interaction (27). This longer CDR loop displays conformational changes upon antigen binding. In both native and complexed structures, the binding of anti-peptide Fab 17/9 to various segments from its antigen, influenza virus hemagglutinin, also utilizes a mostly wide flat surface. Comparison of these structures has shown that the CDR H3 loop exhibits a large conformational change upon antigen binding resulting in two very distinct structures for the bound and free H3 loop (28).

In contrast to these antibodies, the analogous region in the MAK33 structure is slightly convex with a shallow groove or channel running the width of the molecule resulting in a bipartite binding surface. These structural differences suggest a possible alternate recognition mode for MAK33 consistent with its requirement for the dimeric form of the antigen. The longer CDR H3 loop contains 10 residues and, as seen in the lysozyme antigen/antibody complexes, would be expected to affect the surface presented for antigen binding. Although the corresponding electron density for this loop is weak in our structure, thus prohibiting the determination of its specific conformation, this disorder is suggestive of a possible direct interaction with the antigen such that upon binding this loop might assume a stable conformation.

The molecular surfaces presented for complex formation between the light and heavy chain of any given antibody molecule comprises symmetrical arrangements of the β-strands of each individual domain (29). As shown in Fig. 2, the intersubunit contact surfaces of variable and constant domains differ despite the overall similarity in domain structure. In the MAK33 structure the variable domains interact along the edge of the sheet made up of β-strands HGCD such that the contact residues extend along βC and βH and include those of the adjacent loops. The constant domains, however, wrap around each other and interact face to face such that their respective β sheets comprising strands DEBA are at an angle of roughly 60 degrees. Direct contact extends along all these β-strands. These differential interactions are likely to influence the folding pathway of the intact antibody and might be related to differences in sequential ordering or rate-limiting effects of certain steps as compared for single domains or multidomain chains.

Proline Environments: Implications for Folding Studies—Antibody molecules in general have a high degree of primary
sequence conservation. In particular, they contain a number of conserved proline residues in both the light and heavy chain domains. These residues appear to play important structural roles by stabilizing the common scaffolding motifs that maintain the characteristic Ig domain secondary structure and also factor in the determination of the orientations of the hinge regions between individual domains. As observed in previous antibody structures, within the MAK33 Fab several of these conserved proline residues adopt the cis conformation. In general, the locations of the cis prolines can be broadly divided into those that are located near chain termini appearing on the molecular surface and those that are buried within the domain interfaces. The final assembly of the native Ig molecule must include steps to facilitate the isomerization process to the thermodynamically less stable cis form. Therefore, examination of the environment of these cis prolines within the context of the entire antibody molecule helps to differentiate them in terms of any potential roles they might play in protein folding or assembly, such as reactivity with respect to PPIases, the FK506 binding protein, or cyclophilin (2, 30, 31).

In the MAK33 structure a conserved cis proline residue is found at position Pro-L9. This proline is in the N-terminal β segment of the light chain and is generally exposed to the solvent with a solvent-accessible surface of 18 Å². The cis conformation appears to be stabilized by an additional H-bond to Oγ1 of Thr-L103 on an adjacent β strand (Fig. 3A). The equivalent position in the heavy chain is occupied by a glycine residue that is also important in maintaining intradomain contacts. However, near the C terminus of the heavy chain there is a somewhat analogous cis proline residue at Pro-H191. This residue is also exposed to the solvent, with an accessible surface area of 17 Å², and would appear to be able to readily interact with potential isomerases. The cis conformation of this

**FIG. 3. Stereo models of cis-proline environments.** Oxygen atoms are red, carbon atoms are gray, and nitrogen atoms are blue. Subsets of hydrogen bond distances are shown by dotted lines. A, cis-ProL9; B, cis-ProH191; C, cis-ProL142; D, cis-ProH149 and cis-ProH151.

---

2 Nomenclature for sequences: L, light chain; H, heavy chain; e.g. Pro-L9.
residue is most likely stabilized by stacking on the preceding residue Trp H190 (Fig. 3B).

The only other cis proline within the light chain is found at position Pro-L142. This conserved residue is located in the loop connecting β-strands B and C of the constant domain and extends into the elbow region. The cis isomer is stabilized by an H-bond from the carbonyl to the neighboring His-L199. In addition, further stability is afforded by the improved stacking arrangement on the residue immediately preceding it, Tyr-L141 (Fig. 3C). This cis proline residue has an accessible surface of only 10 Å², suggesting that interaction with a PPIase is restricted once the light chain has assumed its tertiary fold. This highly conserved cis proline is critical for maintaining the intrastrand domain orientation.

The two final cis proline residues, His-149 and His-151, are located in close proximity within the heavy chain. It has been previously suggested that the trans to cis isomerization of Pro-H151 is the rate-limiting step in the overall folding process of the MAK33 Fab (2). The present structure demonstrates that this cis isomer is stabilized by the structural constraints of its immediate neighborhood (Fig. 3D). This residue is located in the upper edge of the constant domain of the heavy chain, where it protrudes into the elbow region and helps define the intrastrand angle between the VH and VC11 domains. It is in an analogous location to that of Pro-L142 with respect to the secondary structure arrangement of the individual domains of each protein chain. However, the pairings of heavy and light chains occur in such a way that Pro-L142 is not directly involved in the interstrand constant domain contact site but is within ~4–5 Å of the terminal residues (His-111 through His-113) of the heavy chain variable domain. The Pro-H151 residue, in the context of the heavy chain alone, has a large solvent accessible surface area of 23 Å², but in the intact Fab it is positioned within the constant domain subunit interface. These structural observations provide an explanation for the differential interplay of isomerization and association reactions of individual domains versus longer fragments (6).

However, it is clear from this structure that a critical aspect of the stabilization of Pro-H151 comes from the presence of the other cis proline just 2 residues removed, at position H-149. This proline residue has a very small accessible surface area, 3 Å², and is involved in numerous contacts. These include, stacking on Phe-H148 and the direct involvement in an intrachain contact with the His-H201 side chain (Fig. 3D). Pro-H149 is also buried within a hydrophobic pocket formed by Leu-H8, Ala-H203, Thr-H123, and Pro-H151. This stacking pattern is observed in the closely related antibody, 3D6 (25), as well as where, despite changes in individual residues, the overall structural interactions of this immediate region are conserved.

Binding of MAK33 Fab Peptides to BiP—BiP has for some time been an appealing target of studies to understand how chaperones identify denatured proteins. Antibodies are the best-studied natural substrate of this chaperone (7, 10, 12–15). Potential binding sites identified within the variable and constant domains of the light chain, designated LM1–176, (scores ranging from 6 to 10 and higher) and another eight peptides corresponding to potential binding sites within the light chain of MAK33, designated HM1–166, (scores ranging from +6 to +20) for their ability to bind to BiP and stimulate the ATPase activity of BiP in vitro. One additional peptide (HM164) was tested despite its strongly negative BiP score, because of its particular location within the constant domain of the heavy chain (as discussed further below).

The sequences and BiP scores of these peptides are presented together with the results of the ATPase stimulation in Table II. Of the 17 peptides tested, two peptides, one each within the variable and constant domains of the light chain, and three peptides within the heavy chain stimulated the ATPase activity by factors ranging from 2.0 to 3.0, similar to
Synthetic peptides are named according to the sequence number, within the antibody chain, of the first residue of each heptapeptide. Those listed as LM and HM correspond to sequences in the MAK33 light and heavy chains, respectively. Each BiP score is the sum of the individual scores for the residues present at positions 1–7, which were calculated using the BiP score program described by Knarr et al. (11). The stimulation factor of the ATPase activity of BiP was determined for each peptide as described under “Experimental Procedures.” The values shown are averages of at least three experiments.

| Peptide | Sequence | Score for position | BiP score | ATPase stimulation |
|---------|----------|--------------------|-----------|-------------------|
| LM1     | MIVLVTQ  | 4  1  0  2  3  5 +11 | 1.3 ± 0.2 |
| LM72    | FTSINS   | 7  2  4  0  2  0 +13 | 1.7 ± 0.2 |
| LM84    | FGMYFQCQ | 7  2  0  3  2  5 +19 | 1.1 ± 0.2 |
| LM91    | QSNPSWL | 1  2  6  4  12 5 +14 | 1.6 ± 0.2 |
| LM93    | NSWPLTF | 0  2  6  1  3  2 +12 | 2.5 ± 0.3 |
| LM134   | VCFNLNF  | 0  0  8  1  0  0 +11 | 1.3 ± 0.1 |
| LM161   | LNSWTDIQ | 2  0  4  0  5  5 +11 | 1.2 ± 0.3 |
| LM176   | MSSTTL   | 4  2  0  2  3  2 +10 | 2.9 ± 0.2 |
| HM14    | PGGSLKL  | 1  2  0  2  3  2 +9  | 2.0 ± 0.4 |
| HM22    | SGFTFSD  | 3  2  8  2  3  2 +6  | 1.0 ± 0.3 |
| HM24    | FTFSDYY  | 7  2  8  2  0  1 +20 | 1.4 ± 0.1 |
| HM31    | MYWVRQT  | 4  1  6  2  1  0 +9  | 2.3 ± 0.2 |
| HM33    | WVRQFTE  | 5  1  2  0  0  0 +8  | 1.1 ± 0.1 |
| HM106   | AMDYAGWY | 5  3  2  3  4  1 5 +7  | 1.2 ± 0.1 |
| HM137   | MYTLGLC  | 4  1  1  1  2  0 +8  | 1.1 ± 0.1 |
| HM164   | GVHTFFPA | 2  1  0  2  0  12 | 1.0 ± 0.2 |
| HM166   | HTPFVAVL | 0  2  8  1  1  3 5 +12 | 3.0 ± 0.4 |

*These values were taken from Knarr et al. (11).

Values reported previously for other peptides (11). An additional three peptides displayed consistent, but lower stimulatory effects of 1.4 to 1.7. The remaining nine peptides caused no significant stimulation of the ATPase activity and, therefore, were not further considered as likely BiP-binding sites.

Location of the Possible Binding Motifs on the Three-dimensional Structure of MAK33—The BiP score based on primary structure analysis indicates possible BiP binding motifs based on the residue characteristics within short peptide segments. Our structure allows detailed topological analysis of binding motifs within the context of the overall antibody architecture. We observe a strong correlation of high BiP scores with the structural criteria of burying buried within the antibody molecule (Fig. 4). The majority of identified peptides involve residues that participate in interdomain contact interactions within one subunit or between the light and heavy chains, suggesting that these generally hydrophobic sequences would only be exposed during folding and assembly or under denaturing conditions. A more detailed analysis of the location of specific peptides provides additional information on the exact nature of possible interactions.

The conserved C11 domain is known to be an essential target zone in the binding of antibodies by BiP in vivo. Sites were identified within the sequence of MAK33 that are found at the interface of the constant domains. In particular, HM137 and LM134 are located at analogous positions within strand βB of the heavy and light chain constant domains, respectively. These peptides, located within the core of their respective domains and ~9 Å apart, both show moderate BiP scores but very low ATPase stimulation. In contrast, examination of another set of analogous sites (HM166 and LM176) reveals a tendency toward high BiP score and increased ATPase stimulation at interface positions nearer to the intrachain elbow regions. However, the two overlapping peptides located on βD of the heavy chain constant domain, HM164 and HM166, have very different BiP scores and ATPase stimulation values. HM166, which is closer to the elbow region of the heavy chain and the intrachain contact surface, has significantly higher values for both, suggesting that binding affinity in this region is high and that this region might be directly involved with the potential conformational change that is associated with substrate binding by BiP. Moving away from this immediate elbow region by even a few residues (as seen with HM164) would then be consistent with significantly reduced binding affinity as evidenced by the negative BiP score and further supported by the lack of ATPase stimulation. On the other hand, the light chain peptide LM176 (βE) is also located close to its respective elbow region at the interface and is roughly 7 Å from the heavy chain βD strand. This peptide shows values close to those of HM166, and its ATPase stimulation is the highest among the light chain peptides tested.

One additional heavy chain peptide displayed a moderate BiP score as well as a significant ATPase stimulation factor, indicating that it may be a favored binding site. This peptide, HM14, is located on the outer surface of the protein very near the N terminus of the heavy chain. The most likely explanation for binding in this region would be to stabilize the growing protein chain and promote the proper folding of the protein as it is being produced in the ER. The presence of a bound chaperone might be expected to affect positively the ratio of folded to aggregated heavy chain and promote antibody production.

Two sequential light chain peptides, LM91 and LM93, are located on the βH strand of the variable domain and are positioned progressively farther away from the elbow and the core of the domain and closer to the protein surface. LM93 actually comprises part of a CDR loop. Due to the orientation of variable and constant domains, this progression in primary sequence is also leading into the area of the light intrachain domain interface. Because these sequences are at the end of the variable domain, the chaperone binding might be required to stabilize the growing light chain so that its final domain can fold properly. This would be analogous to binding at HM14, as mentioned above, and perhaps more importantly, would also help to orient the growing chain such that the two domains are positioned properly with respect to each other. In contrast, LM81 and LM84 are also located within the elbow region but on the outer side of the light chain V-C junction. These peptides are far from the dimer interface and do not appear to bind BiP.

**DISCUSSION**

The results reported here provide an additional perspective from which to consider questions, previously addressed by solution studies, relating to protein folding and antibody assembly. The detailed structural view of any particular residue(s) of
interest, e.g. Pro-H151, whose isomerization was implicated as the likely rate-limiting step in the folding of the entire Fab, emphasizes not only its own contribution but also the context within which these effects are implemented. From the MAK33 structure it is difficult to discern why the isomerization of His-151 is a more likely candidate for the rate-limiting step than His-149. However, it is clear that the combined effects of these two highly constrained residues make this overall site structurally unique as compared with the immediate environment of the other cis prolines. Perhaps it is the combination of the isomerization of both residues that contribute to the rate-limiting step in the folding pathway as opposed to the isomerization of either of these in particular.

As has been suggested previously, it is plausible that being temporarily slowed by cis/trans proline isomerizations is not necessarily a disadvantage among refolding procedures but could possibly provide the time needed to allow previously trapped refolding steps to correct or complete themselves (36, 37). From biochemical studies it is known that individual light chains are able to be completely reversibly denatured (3), whereas similar treatment of heavy chains results in significant losses due to aggregation.4 This might be related to the differential reactivity of prolines in these two components, as the ease at which these prolines assume their final conformation might be related to the degree to which exposed B-strands are involved in nonspecific aggregate formation. The surface characteristics of individual domains would affect interchain contacts as well as possible interactions with any other proteins necessary for proper antibody folding, such as prolyl isomerases and chaperones.

In certain myeloma cell lines it is the variable domain of the light chain that is specifically associated with BiP (38). This increased binding has been attributed to the fact that the variable domains exhibit a greater degree of sequence variability, especially relating to their function in antigen recognition, and hence might be harder to fold than constant regions. As a result improperly exposed hydrophobic surfaces would become likely binding sites for chaperone interaction.

The primary sequence of a peptide segment does influence its general propensity to form fundamental secondary structure elements. The alternating hydrophobic nature of those segments identified by the BiP score suggests that better binders are involved in β-strand formation. However, structural comparison of potential BiP-binding sites within MAK33 also emphasizes that, in addition to the presence of an individual secondary structure feature, it is the protein environment with respect to intra- and interdomain orientations that is critical for recognition. The overall three-dimensional structure of such a recognition site is determined by key characteristics conserved among proteins of a given class despite differences in the details of individual examples. In addition, the characteristic manner of pairing in antibodies of variable versus constant domains is significantly different in general, and therefore, the interaction with BiP may be influenced by the different quaternary organization as well. Because interchain contacts are in fact essential for the pairwise domain orientations, binding of BiP to “hot spots” within such contact regions might be expected to compete with the proper antibody subunit interactions needed to complete the native molecule. The inability of the heavy chains to fold, as observed in vitro, may be compensated for by chaperone binding in vivo.

Taken together, these findings suggest that, as an ER-specific hsp70 class chaperone, BiP might have a dual function. Interactions with many misfolded proteins could take place through binding of short hydrophobic sequences in a manner much like that described for DnaK. The apparent ambiguity in the scores for antibody peptides, i.e. that not all high scoring potential binders are located at the domain interfaces, is consistent with BiP’s purely primary sequence-specific mode of interaction. A more specific functional role of BiP, however, might utilize an interaction analogous to that seen in the complexes of bacterial immunoglobulin type chaperones involved in pili assembly with their specific cellular targets. In contrast to the sequence-based binding of exposed residues, the crystal structures of these complexes reveal another mode of interaction that is based on the swapping of a β-strand between chaperone and target (39, 40). The crystal structure of the peptide binding domain of the BiP homologue, DnaK, suggests that in both proteins this domain is largely β-structured (41). Although this β domain has a topology that is different from that in a typical immunoglobulin fold, the presence of β sheet motifs in both partners suggests that in this case the heavy chain binding mode of BiP may utilize a β-strand donor or complementarity as seen with PapD and FimC. Thus this more specific function of assisting antibody assembly might utilize the structural requirements of the pairing of two β domains.

The overall folding pathway of antibodies involves a series of interactions not only between the protein subunits but also with folding factors such as isomerases or chaperones. It is therefore likely that there is a kinetic competition for binding between these proteins that must dictate the rules of this complicated interplay. This is also consistent with the findings that, despite the highly conserved secondary structure, similar residues within different individual domains can have quite diverse influences on the formation of the native folded structure. It seems that antibodies have evolved to modulate a highly conserved structural motif (the immunoglobulin fold) according to different functions (antigen binding, complement binding, stability, association) through intrinsic signals for β sheet and dimer formation as well as extrinsic signals, i.e. for interaction with BiP. The three-dimensional structure of MAK 33 Fab provides an important starting point for further analyzing these signals by a systematic structure-based mutagenesis approach.

Acknowledgments—We thank Alfred Engel for sequencing the MAK33 cDNA, Susanne Modrow for the synthesis of peptides, Hauke Lilie for helpful discussions and unpublished data, Mary-Jane Gething for insightful discussions of the BiP score, Helmut Lenz for sharing his knowledge on immunoglobulins throughout this work, and Juia-huai Wang for helpful discussions and critical reading of the manuscript.

REFERENCES
1. Buckel, P., Hübner-Parajsz, C., Mattis, R., Lenz, H., Haug, H., and Beaucamp, K. (1987) Gene 51, 13–19
2. Lilie, H., Lang, K., Rudolph, R., and Buchner, J. (1993) Science 2, 1490–1495
3. Lilie, H., McLaughlin, S., Freedman, R., and Buchner, J. (1994) J. Biol. Chem. 269, 14290–14296
4. Lilie, H., Rudolph, R., and Buchner, J. (1995) J. Mol. Biol. 248, 190–201
5. Lilie, H., and Buchner, J. (1995) FEBS Lett. 362, 43–46
6. Thies, M. J. W., Mayer, J., Augustine, J. G., Frederick, C. A., Lilie, H., and Buchner, J. (1999) J. Mol. Biol. 293, 67–79
7. Hendershot, L., Bole, D., Kohler, G., and Kearney, J. F. (1987) Cell Biol. 104, 761–767
8. Bole, D. G., Hendershot, L. M., and Kearney, J. F. (1986) J. Cell Biol. 102, 1558–1566
9. Munro, S., and Pelham, H. R. (1986) Cell 46, 291–300
10. Blund-Elguindi, S., Cwirla, S. E., Dower, W. J., Lipshutz, R. J., Sprang, S. R., Sambrook, J. F., and Gehring, M.-J. H. (1995) Cell 75, 717–728
11. Knarr, G., Gehring, M.-J. H., Modrow, S., and Buchner, J. (1995) J. Biol. Chem. 270, 25789–25794
12. Knarr, G., Modrow, S., Todd, A., Gehring, M.-J. H., and Buchner, J. (1999) J. Biol. Chem. 274, 29850–29857
13. Schmidt, M., and Buchner, J. (1992) J. Biol. Chem. 267, 16829–16833
14. Collaborative Computational Project Number 4 (1994) Acta Crystallogr. Sect. D. Biol. Crystallogr. 50, 760–763
15. Otwinowski, Z. (1989) Data Collection and Processing. Proceedings of the CCP4 Study Weekend Daresbury, UK
16. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
17. Rini, J. M., Schulze-Gahmen, U., and Wilson, I. A. (1992) Science 255, 959–965

4 H. Lilie and J. Buchner, unpublished results.
18. Navaza, J. (1994) Acta Crystallogr. Sect. A 50, 157–163
19. Brünger, A. T. (1992) A System for X-ray Crystallography and NMR: X-PLOR, Version 3.1, Yale University Press, New Haven, CT
20. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
21. Laskowski, R. A., McArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Cryst. 26, 283–291
22. Wiech, H., Buchner, J., Zimmermann, M., Zimmermann, R., and Jakob, U. (1993) J. Biol. Chem. 268, 7414–7421
23. Atherton, E., Fox, H., Logan, C. J., Harkiss, D., Sheppard, R. C., and Williams, S. B. J. (1978) Chem. Soc. Chem. Commun. 13, 537–539
24. Chothia, C., and Lesk, A. M. (1987) J. Mol. Biol. 196, 901–917
25. He, X. M., Ruker, F., Casale, E., and Carter, D. C. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 7154–7158
26. Padlan, E. A., Silverton, E. W., Sheriff, S., Cohen, G. H., Smith-Gill, S. J., and Davies, D. R. (1998) Proc. Natl. Acad. Sci. U. S. A. 86, 5938–5942
27. Lescar, L., Pellegrini, M., Souchon, H., Tello, D., Poljak, R. J., Peterson, N., Greene, M., and Alzari, P. (1995) J. Biol. Chem. 270, 18067–18076
28. Schulze-Gahmen, U., Rini, J. M., and Wilson, I. A. (1993) J. Mol. Biol. 243, 1098–1118
29. Branden, C., and Tooze, J. (1991) Introduction to Protein Structure, Garland Publishing, Inc., New York
30. Lang, K., Fischer, G., and Schmid, F. X. (1987) Nature 329, 268–270
31. Harrison, R. K., and Stein, R. L. (1990) Biochemistry 29, 3815–3816
32. Flynn, G. C., Chappell, T. G., and Rothman, J. E. (1989) Science 245, 385–390
33. Flynn, G. C., Pohl, J., Flocco, M. T., and Rothman, J. E. (1991) Nature 353, 726–730
34. Knittler, M. R., and Haas, I. G. (1992) EMBO J. 11, 1573–1581
35. Kaloff, C. R., and Haas, I. G. (1995) Immunity 2, 629–637
36. Pitzsyn, O. B., and Uversky, V. N. (1994) FEBS Lett. 341, 15–18
37. Pitzsyn, O. B. (1996) Nat. Struct. Biol. 3, 488–490
38. Skowronek, M. H., Hendershot, L. M., and Haas, I. G. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 1574–1578
39. Sauer, F. G., Futterer, K., Pinkner, J. S., Dodson, K. W., Hultgren, S. J., and Waksman, G. (1999) Science 285, 1058–1061
40. Choudhury, D., Thompson, A., Stojanoff, V., Langerman, S., Pinkner, J., Hultgren, S. J., and Knight, S. D. (1999) Science 285, 1061–1066
41. Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. (1996) Science 272, 1606–1614
42. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
43. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Proteins Struct. Funct. Genet. 11, 281–296
