PROTEINS: Structure, Function, and Genetics 42:269–278 (2001)

Structure of an Intermediate in the Unfolding of Creatine Kinase

Timothy I. Webb and Glenn E. Morris*
MRIC Biochemistry Group, North East Wales Institute, Wrexham, United Kingdom

ABSTRACT **The homodimeric muscle isoform of creatine kinase (MM-CK) unfolds on exposure to low levels of guanidinium chloride (GdmCl) to yield a partly folded monomeric intermediate. Those regions of MM-CK that experience local unfolding were previously identified through an extensive study of antibody accessibility and protease sensitivity. Since these studies were completed, the coordinates of the rabbit isoform (MM-CK) were released. In light of this, we have determined the minimum changes to this structure required to explain our data on protease and epitope accessibility in the intermediate. We propose that the observed changes occur through (a) disruption of the monomer-monomer interface during dissociation, (b) separation and/ or unfolding of domains or subdomains, and (c) the partial unfolding of solvent-exposed helices. The proposed structure for the intermediate is consistent both with current models of unfolding intermediates and the results of independent studies pertaining to the unfolding of creatine kinase. Proteins 2001;42:269–278.**

Key words: protein folding; domain structure; antibody; protease; epitope; guanidinium; molten globule; dimer

INTRODUCTION

GdmCl-induced unfolding of MM-CK at 25°C does not conform to models of a two-state transition.[1–4] A plateau in the λ_{max} and far-UV CD unfolding profiles between 0.8 M and 1.2 M GdmCl signifies the population of an intermediate.[2–4] A maximum in the enhancement of ANS fluorescence is also observed within this same concentration range.[2] It has been concluded that an initial unfolding transition results in the population of a stable unfolding intermediate with the characteristics of the molten globule.[2–4] Additional experiments indicate that this transition is coincident with dissociation and expansion of the isolated subunit.[2,4]

Most previous studies concerning the structure of the unfolding intermediate of MM-CK have relied on physical and chemical techniques. Where such techniques provide data on local unfolding, they have relied on a few specific residues, for example W211, C74, and C146.[2,5,6] As a consequence, the results of these studies do not provide a comprehensive account of the disruption of local structure in the unfolding intermediate. If such data were available, it may be possible to assess the relevance of various hypothetical models for unfolding intermediates. In the equilibrium unfolding intermediate of a dimeric, multidomain protein like MM-CK, for instance, local unfolding may arise as a result of the disruption of subunit structure after dissociation, the sequential unfolding of domains, or hierarchic unfolding within individual subunits. These models are not necessarily mutually exclusive.

One experimental approach that may provide the necessary data is use of a panel of proteases (and monoclonal antibodies) to probe local unfolding in the intermediate. The enhanced susceptibility of an intermediate to proteolysis is primarily a result of unfolding within the hydrophilic "shell."[7–10] Moreover, if complete degradation is checked (and the products of proteolysis accumulate), the nonpolar core must remain structured and relatively rigid. The main determinant of this "limited" proteolysis is the ability of a local segment to unfold in the absence of a global perturbation of protein structure.[10,11] The unfolding is necessary for the protein substrate to adopt a particular local conformation before docking with the protease and subsequent cleavage.[11] This process requires a significant disruption of native backbone conformations in a segment that includes (at least) six residues on each side of the scissile bond. The correctness of using proteolysis to identify the stable core of a protein under mild denaturing conditions was previously verified.[9] For example, although native holo myoglobin is resistant to proteolysis, helix F, which coordinates the haem, is substantially digested in apomyoglobin.[9] Similarly, a stable nicked protein was isolated after excision with pepsin of an unfolded β-sheet from the A-state (pH 2.0/0.1 M NaCl) of α-lactalbumin.[10] These results were in agreement with those obtained by using the usual biophysical techniques.[9,12,13]

A study of the accessibility of the chick MM-CK unfolding intermediate to antibody binding and proteolysis provided early evidence for the sequential unfolding of domains.[14,15] It was proposed that, although residues M1–E166 constitute an unfolded N-terminal domain, residues G167–K381 constitute a relatively stable C-terminal domain. However, subsequent publication of the crystal structures of homologous chick Mi_b-CK and rabbit MM-CK

*Correspondence to: Glenn E. Morris, MRIC Biochemistry Group, North East Wales Institute, Wrexham LL11 2AW, UK. E-mail: morrisge@newi.ac.uk

Received 1 June 2000; Accepted 28 September 2000

indicated that the N-terminal domain comprises (approximately) only the first 100 residues.[16,17] The local unfolding associated with a protease susceptible region in the middle of the M-CK subunit (R135–E166) remains unexplained. The release of the coordinates for the structure of MM-CK has made it possible to reassess earlier unfolding data. We now show that much of the local unfolding identified in the intermediate occurs as a result of a disruption of the monomer-monomer interface during dissociation. However, although each subunit maintains native-like structure within the two subdomains that comprise the C-terminal domain, there is a significant disruption of structure within the N-terminal domain.

MATERIALS AND METHODS

The protein modeling was performed on a Silicon Graphics, Inc. workstation by using INSIGHT II (Molecular Simulations Incorporated). The coordinates of the crystal structure of MM-CK (2CRK.PDB) and Mi_b-CK (1CRK.PDB) are available from the Brookhaven database (http://www.pdb.bnl.gov). The residue contacts were identified on-line by using the CSU software.[18] The accessible molecular surface calculations were performed on-line (http://swift.embl-heidelberg.de/servers2) by using routines implemented in the WHATIF program. The probe radius used was 1.4 Å. Additional routines within WHATIF were used to identify the hydrogen bonds. In the potential hydrogen bond calculation, parameters used were as follows: a maximal donor atom-acceptor atom distance of 3.5 Å, a maximal hydrogen atom-acceptor atom distance of 2.5 Å, a maximal donor atom-hydrogen atom-acceptor atom angle of 60°, a maximal hydrogen atom-acceptor atom-antecedent acceptor atom angle of 90°. The optimized network of hydrogen bonds was also identified.[19] The monomer-monomer interface was analyzed on-line by using the Protein-Protein Interaction Server (http://www.biochem.ucl.ac.uk/bsm/PP/server). Finally, the hydrophobic folding units (of Mi_b-CK) were identified by using a routine implemented on-line (http://protein3d.ncifcrf.gov:1025/tsai/).[20]

RESULTS AND DISCUSSION

We previously used protein microsequencing to identify the principal trypsin, chymotrypsin, and endoproteinase Glu-C digestion products that become exposed after treatment of chick MM-CK with 0.63 M GdmCl at 37°C.[21] The 11 major digestion sites and 7 minor digestion sites identified are listed in Table I (also Fig. 1a). Table I also shows the segregation of these digestion sites within five protease susceptible regions. The minor digestion sites are classified into two that occur early in the time-course and five that occur later. The results of this proteolysis experiment were interpreted as evidence for the population of an equilibrium unfolding intermediate under the prevailing conditions. An additional minor digestion site was identified during proteolysis of a kinetic folding intermediate trapped at low temperature (Table I). However, both experimental approaches yielded very similar digestion patterns. A number of antibodies that recognize epitopes revealed in these intermediates have also been identified.[14,22,23] In the present study, we use these digestion sites and epitopes to identify the local unfolding that occurs during transition of native MM-CK to the unfolding intermediate. Although our model of the unfolding intermediate is based on the structure of rabbit MM-CK,[17] the chick M-CK and rabbit M-CK sequences (Fig. 1a) show an identity of >90%.[24] We propose that local unfolding occurs as a result of three major changes:

TABLE I. Digestion Sites

Region	Target
I	
T	K32
V	E37
C	Y39
II	
C	F68
V	E80
III	
T	R135
C	Y140
T	R148
T	R151
T	R152
V	E166
IV	
C	L176
V	E181
V	E183
C	L201
[a]T	R209
[a]T	R215
V	
V	E262
[b]V	E275

The digestion sites are classified into five protease susceptible regions. While major digestion sites are confined to protease susceptible regions I, II, and III, minor digestion sites are confined to protease susceptible regions IV and V. The protease (endoproteinase Glu-C, chymotrypsin, trypsin) responsible is also indicated (V, C, T). Two minor digestion sites were observed early in the time-course.[a] Another digestion site was identified in the kinetic folding intermediate.[b]

(a) disruption of the monomer-monomer interface during dissociation,
(b) separation and/or unfolding of domains or subdomains, and
(c) partial unfolding of solvent-exposed helices.

In the analysis that follows, we shall show in detail how all the proteolytic digestion sites and antibody accessibility changes can be accounted for by the three types of unfolding proposed.

Disruption of the Monomer-Monomer Interface During Dissociation

Evidence of disruption at the monomer-monomer interface of MM-CK is provided by six major digestion sites (E37, Y39, Y140, R148, R151, and R152) and three minor digestion sites (L176, R209, and R215). The situation of these digestion sites close to the interface is shown in

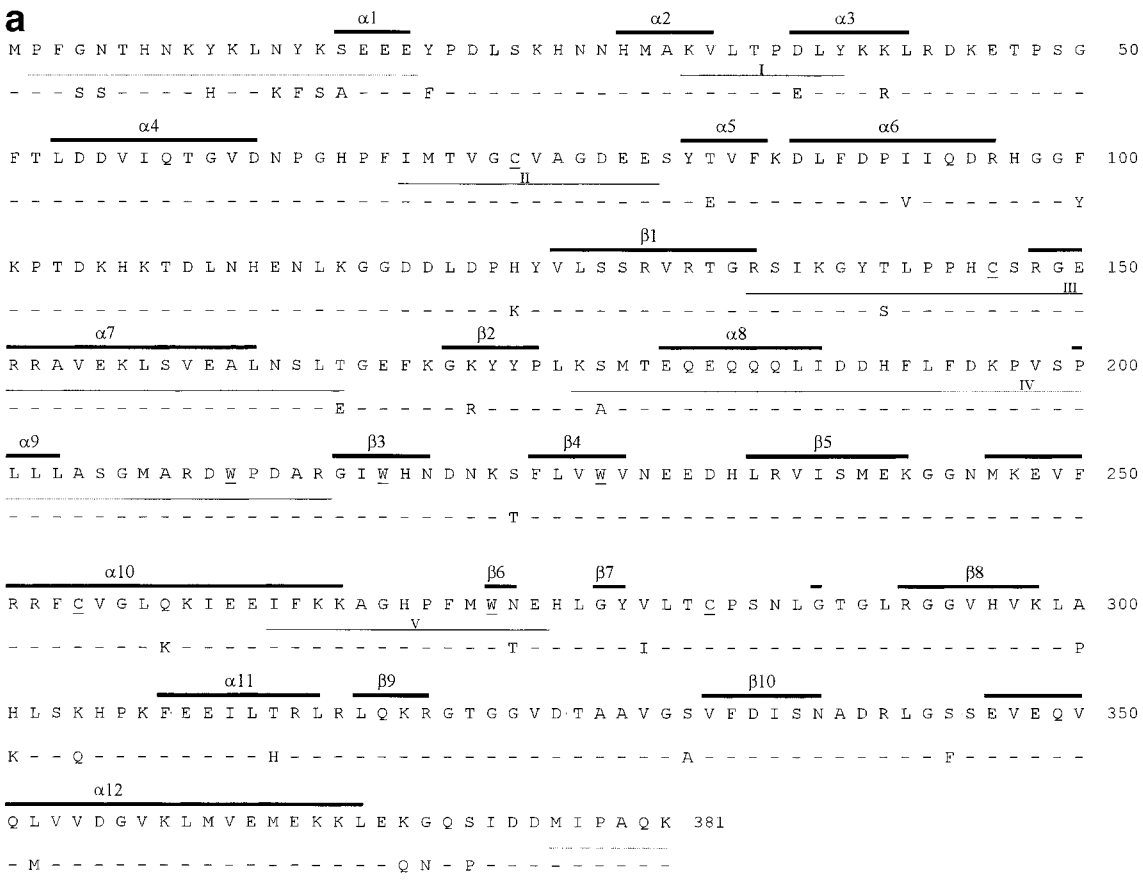

Fig. 1. Sequence and secondary structure of rabbit M-CK. **a:** The sequence of rabbit M-CK is indicated (top). The heavy bars show the positions of α-helices and β-strands in the crystal structure of native MM-CK.[17] Also indicated is residue G288. In Mi_b-CK,[16] this residue comprises the third element of a small β-sheet (β6, β7, and G288). The sequence of chick M-CK is indicated for comparison (bottom). However, only residues that differ between the two sequences are shown. Our proteolysis and antibody-binding studies provide evidence for the exposure of residues in the unfolding intermediate of chick MM-CK. The digestion sites, grouped into five (I, II, III, IV, and V) protease susceptible regions (light bars), and the epitopes of three monoclonal antibodies (dashed), are indicated. A number of additional residues referred to in the text are also indicated (underlined). **b:** Each rabbit M-CK subunit is depicted by a topological diagram. α-helices are represented by bars and β-strands by arrows. Each subunit is composed of three hydrophobic folding units (N, C1, and C2). The N-terminal structural domain corresponds to N, whereas the C-terminal structural domain corresponds to C1/C2.

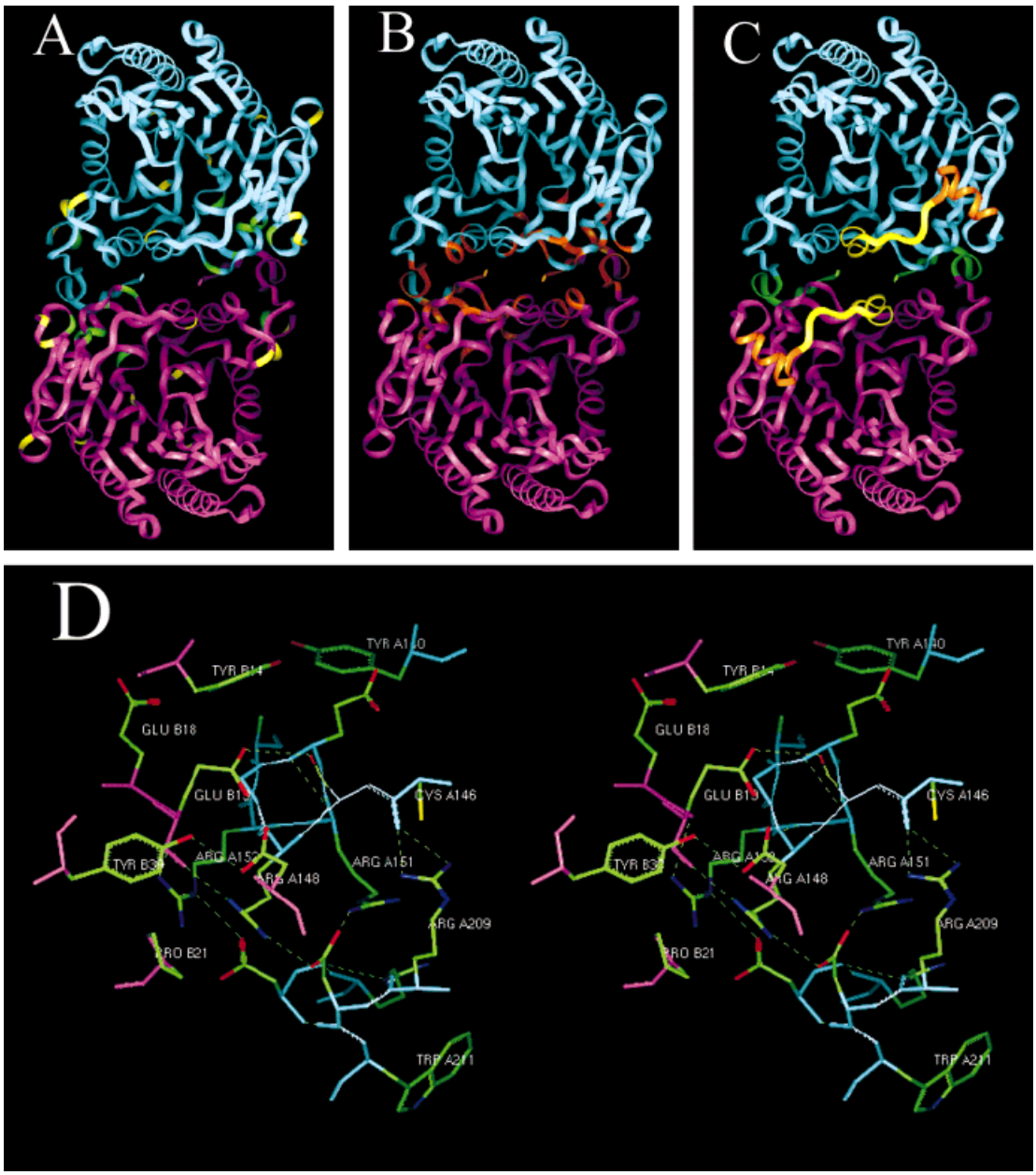

Fig. 2. The monomer-monomer interface of the rabbit MM-CK dimer. **a:** The target residues of all major and minor digestion sites revealed in the unfolding intermediate of chick MM-CK are highlighted. The eight digestion sites situated at the monomer-monomer interface (green) are distinguished from the remaining digestion sites (yellow). A major digestion site (E37) situated adjacent to one of these residues (Y39) is also included. Chain A and chain B of the MM-CK dimer are colored cyan and magenta respectively. **b:** The residues that comprise the actual monomer-monomer interface are highlighted (red-orange). There was a significant increase in the accessible molecular surface of the indicated residues after simple dissociation. **c:** The epitopes of three monoclonal antibodies (CK-STAR, CK-JIL, and CK-JAC) are revealed at the monomer-monomer interface. The epitope of CK-STAR is situated within residues P1–E19 (green). The epitope of CK-JIL is situated within residues F194–G206 (yellow). Although the epitope of CK-JAC overlaps that of CK-JIL, additional residues (E183–L193) are needed for effective binding (orange). **d:** Stereo view of the amino terminus of α7 (along helical axis). Some important interactions between this "hotspot" (S147–R152) and residues (Y14, E18, E19, P21, Y39, Y140, C146, R209, D210, W211, D213, and R215) situated elsewhere at (or close to) the monomer-monomer interface are illustrated. Dashed lines show the presence of hydrogen bonds. The backbone atoms of residues from each subunit are colored to aid identification.

Figure 2a. A comparison of Figure 2a with the actual monomer-monomer interface (Fig. 2b) shows the good correlation between the indicated digestion sites and residues that comprise the monomer-monomer interface. The residues of this interface may be grouped into six interface residue segments (IRS; Table II). The indicated digestion sites provide evidence that three of the six segments (IRS3, IRS4, and IRS6) are disrupted in the unfolding intermediate and exposed on the surface. In addition, another two segments (Fig. 2c), IRS1 and IRS5 (see later), are recognized by monoclonal antibodies specific for denatured CK.[22,23] We conclude that the transi-

TABLE II. Interface Residue Segments

IRS	Interface residues
1	ASN 8, LYS 9, TYR 10, LYS 11, TYR 14, GLU 17, GLU 18, GLU 19, TYR 20, PRO 21, ASP 22
2	PRO 48, SER 49, GLY 50, ASP 54, ILE 57, GLN 58, VAL 61, ASP 62, ASN 63
3(a)	TYR 140
3(b)	HIS 145
3(c)	SER 147, ARG 148, GLY 149, GLU 150, ARG 152, ALA 153
4	LYS 177
5	LYS 196
6	LEU 203, ALA 208, ARG 209, ASP 210, TRP 211, ASP 213

The monomer–monomer interface was identified using a routine implemented on the Protein–Protein Interaction Server. The definition of the monomer–monomer interface involves the identification of residues for which there is a decrease in Accessible Surface Area after subunit association.[25] Interface residues separated by more than five residues are allocated to different IRS. In Mi_b-CK, due to a mutation in the consensus sequence, the sidechain of A140 (H145 of MM-CK) is not situated at the monomer–monomer interface. In consequence, the routine partitions IRS3 to produce two separate IRS (3 (a) and 3 (c)).

tion from native MM-CK to the unfolding intermediate is characterized both by the exposure of the monomer-monomer interface and by the disruption of local structure in the IRS that comprise this interface.

In a study of protein dimers, the intermolecular interactions of particular IRS ("hotspots") were observed to dominate the association.[25] We have analyzed the structure of rabbit MM-CK in terms of the contribution of individual IRS to the accessible molecular surface revealed at the monomer-monomer interface. According to this criteria, interactions involving the main-chain and side-chain atoms of residues within IRS1 and IRS3 dominate. Indeed, IRS1 and IRS3 account for 33% and 34% of this surface, respectively. Furthermore, in agreement with earlier observations,[25] the side-chain atoms of residues within the dominant IRS account for a large proportion (50%) of the rabbit MM-CK interface. There is also a correlation between the accessible molecular surface contributions of an IRS and the distribution of hydrogen bonds.[25] The 20 intermolecular hydrogen bonds that connect the two subunits of MM-CK are listed in Table III. Fourteen of these intermolecular hydrogen bonds involve residues (S147–R152) within IRS3. The same interactions are observed within the structure of an equivalent dimer of Mi_b-CK.[26] However, the residues involved, S142–R147 of Mi_b-CK, are included within a separate IRS (Table II). The conservation of this network of intermolecular hydrogen bonds (and an ion pair) centered around residues S147–R152 suggests that the network may be an integral component of dimer structure. We propose that residues S147–A153 of IRS3 constitute a hotspot.

In native MM-CK, the hotspot corresponds to the N-cap (S147) and N-terminus of helix α7 (Fig. 1a). The principal interactions that occur in native MM-CK between residues at the N-terminus of α7 and residues elsewhere at the monomer-monomer interface are illustrated in Figure 2d. The side-chain atoms of four of the residues involved account for 30% (Y14 [5.8%], R148 [11.7%], R152 [6%], R209 [5.6%]) of the accessible molecular surface revealed at the interface. This is consistent with the observation that the side chains of a few ARG and TYR residues generally play a dominant role in association.[25,27] Significantly, the target residues of a total of seven (Y39, Y140, R148, R151, R152, R209, and R215) digestion sites are directly involved in interactions between the N-terminus of α7 and residues elsewhere at the monomer-monomer interface. The results indicate that interactions involving the N-terminus of α7, including a network of hydrogen bonds, are completely disrupted in the unfolding intermediate and that local structure around many of the interacting residues is disrupted.

TABLE III. Intermolecular Interactions in MM-CK

Acceptor	IRS	Donor	IRS	Interaction
GLU 18 O	1	ARG 152 NH1	3	hydrogen bond
GLU 19 OE2	1	SER 147 OG	3	hydrogen bond
GLU 19 O	1	ARG 148 NH1	3	hydrogen bond
GLU 19 OE1	1	GLY 149 N	3	[a]hydrogen bond
TYR 20 O	1	ARG 152 NH1	3	hydrogen bond
TYR 20 O	1	ARG 152 NH2	3	hydrogen bond
ASP 54 OD1	2	ARG 148 NH1	3	hydrogen bond/ ion pair
ASP 210 OD1	6	ASN 58 NE2	2	hydrogen bond
ASP 210 OD1	6	ASP 62 N	2	hydrogen bond
		Optimized Network		
ASN 8 O	1	ARG 209 NH1	6	hydrogen bond

The hydrogen bonds were identified using routines implemented in the WHATIF program. With one exception,[a] all potential interactions have previously been reported.[17] An additional interaction was identified using a routine which optimizes the network of hydrogen bonds. Note: all interactions are duplicated as a consequence of two-fold symmetry. The term IRS refers to the Interface Residue Segment to which an interacting residue belongs.

Antibody-binding studies provide additional evidence for the disruption of local structure within the IRS. A sequential epitope at the N-terminus of each subunit is revealed only after transition of native BB-CK to an unfolding intermediate.[23] The exposure of this epitope provides evidence of unfolding around residues in IRS1 (Y14 and E19) that interact with the hotspot (Fig. 2c). Another sequential epitope, centered around residue Y39, also a major digestion site (Table I), is revealed only after transition of native MM-CK to an unfolding intermediate.[28] Although not situated within an IRS, the side chain of residue Y39 is involved in extensive intramolecular interactions with residues in IRS1 and IRS2. These in-

clude nonpolar interactions involving residues E19, F20, P21, L53, and I57, aromatic interactions involving residue F20, and a hydrogen bond to the side chain of residue D54. Furthermore, the side chains of residue Y39 and a residue within the hotspot (R148) may be involved in a weak intermolecular hydrogen bond. Indeed, in an equivalent dimer from Mi_b-CK,[26] the hydrogen bonding potential between the side chains of these residues is fulfilled. Finally, the antibody-binding studies also provide evidence for the exposure of IRS5 (K196) on the surface of the unfolding intermediate. Two antibodies that recognize an MM-CK epitope revealed on denaturation were identified.[22] This epitope was mapped to a conserved segment (F194–G206) that includes IRS5 (Fig. 1a). This epitope (Fig. 2c) is revealed after simple dissociation and mild denaturation of the monomer-monomer interface (unpublished data).

A variety of techniques have been used to follow those changes that accompany the transition from native protein to the (molten globule) unfolding intermediate.[1–6] Significantly, residues identified through residue-specific techniques as exposed in the unfolding intermediate of rabbit MM-CK are also accessible to proteolysis and antibody binding in the unfolding intermediate of chick MM-CK. The exposure of many of these residues is accounted for by unfolding at the monomer-monomer interface. For example, although only one cysteine residue is exposed to chemical modification in each native subunit, additional cysteine residues are revealed in the unfolding intermediate of rabbit MM-CK.[2,4] One of these is likely the same cysteine residue, C146 (Fig. 2d), revealed for chemical modification after dissociation and mild denaturation of rabbit MM-CK at high concentrations of NaCl.[6] A corresponding exposure of C146 in the unfolding intermediate of chick MM-CK is confirmed by the digestion results (Fig. 1a). Indeed, several major digestion sites (Y140, R148, R151, and R152) provide evidence of substantial unfolding around residue C146.

The transition to this rabbit MM-CK unfolding intermediate is also characterized by an increase in the proportion of iodide-quenchable fluorescence and a red shift in λ_{max}.[2] These changes reflect an exposure of the side chain of residue W211 (IRS6), during unfolding at the monomer-monomer interface.[2,4,5] Indeed, similar changes are observed after dissociation and mild denaturation are induced with high concentrations of NaCl.[6] An important role for the indole side chain of this residue in maintaining the quaternary structure of native CK has been shown by using site-directed mutagenesis.[29,30] Again, our proteolysis results confirm the exposure of residue W211 on transition to the unfolding intermediate (Fig. 1a and Table I). Thus, two (early) minor digestion sites (R209 and R215) provide clear evidence of a disruption of local structure within IRS6. Finally, evidence has been presented for the exposure of additional (unknown) tyrosine residues in the unfolding intermediate of rabbit MM-CK.[4] This is consistent with the exposure of the target residues for two major digestion sites (Y39 and Y140) in the chick MM-CK unfolding intermediate (Fig. 1a and Table I). Both these residues are exposed as a result of dissociation and a disruption of the monomer-monomer interface. Moreover, analysis of the crystal structure reveals that there would be a significant increase in the solvent exposure of the side chain of another tyrosine residue, Y14 (IRS1), during simple dissociation (Table II). The antibody-binding studies confirm that IRS1 is revealed in a BB-CK unfolding intermediate.[23]

Separation and/or Unfolding of Domains or Subdomains

Each native subunit of CK is characterized by N-terminal and C-terminal structural domains (Fig. 3a).[16,17] These structural domains also represent functional units.[16,31] The C-terminal domain is capable of autonomous folding after fragmentation.[32] However, the stability (and possibly also the rate of folding) of a fragment that includes the N-terminal domain is influenced by interactions involving the C-terminal domain. There is clear evidence that the N-terminal domain of MM-CK is destabilized in the unfolding intermediate.[14,15,21] The five major digestion sites within protease susceptible region I (K32, E37, and Y39) and protease susceptible region II (F68 and E80) show that the surrounding residues (N27–K45; N63–K86) are unstructured and exposed on the surface of the unfolding intermediate (Table I). As already discussed, antibody-binding studies provide evidence for an additional disruption of structure at the N-terminus of each subunit.[23] We conclude that much of the N-terminal domain (N8–D22, N27–K45, and N63–K86) is significantly disrupted in the unfolding intermediate. In particular, we note that residues (V75–K86) situated at the interface between domains are disrupted and exposed after separation and unfolding of the N-terminal domain (Fig. 3b). The exposure of an adjacent buried cysteine, C74 (Fig. 1a and Fig. 3b), in the unfolding intermediate confirms our own results.[2,4,33]

A likely consequence of the separation of the N-terminal domain is the exposure and destabilization of a complimentary surface within the C-terminal domain. We have identified those residues that show an increase in accessible molecular surface after removal of the N-terminal domain. These residues are arranged within four segments. The first segment (K138, Y140, T141, L142, P144, and H145) coincides with IRS3 and includes two major digestion sites (R135 and Y140). The second segment (M272, W273, N274, H276, L277, V280, L281, T282, C283, P284, S285, and L287) coincides with a region that contains the three small strands (W273–N274; G278–Y279; G288) of a second β-sheet and includes a minor digestion site (E275). Together, the three digestion sites provide evidence both that the two segments are situated at the surface of the unfolding intermediate and that there is a disruption of local structure (Fig. 3c). Our digestion results are confirmed by the observation that unfolding occurs around cysteine residues (C146 and C283) within each of these segments (Fig. 3b).[2,4,34] Two additional minor digestion sites in the C-terminal domain (L201 and R209) provide evidence for the destabilization of a third segment

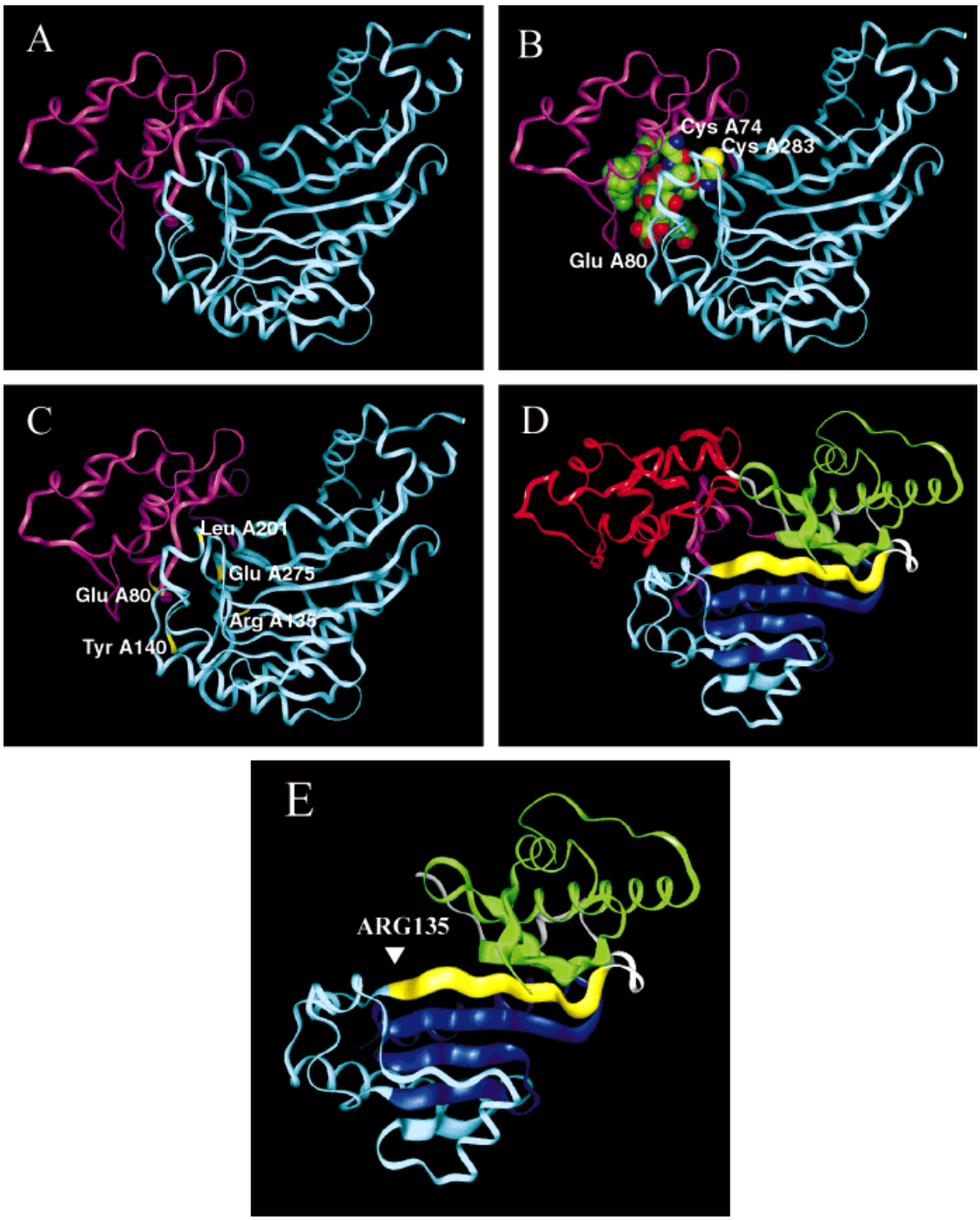

Fig. 3. Domain-domain interfaces in rabbit M-CK. **a:** Each subunit of MM-CK is divided into an N-terminal structural domain, residues 1–105 (magenta), and a C-terminal structural domain, residues 106–381 (light blue). **b:** The situation at the domain interface of residues V75–K86 of rabbit M-CK. Also situated at the domain interface are CYS 74 and CYS 283. It is possible to chemically cross-link these residues across the domain interface.[33] **c:** The situation of digestion sites (yellow) at the domain interface of both domains. **d:** There are two hydrophobic folding units (or subdomains) within the C-terminal domain. The first subdomain (yellow/light blue/dark blue/magenta) and the second subdomain (green) correspond to C1 (S128–G293) and C2 (G294–K381), respectively (Fig. 1b). β1 (V126–R135) and the second β-sheet region (G268–L291) are shown in yellow and magenta, respectively). The third hydrophobic folding unit corresponds to the N-terminal structural domain (red). The N-terminal end of C1 (light blue) includes protease susceptible region III (R135–T166) and protease susceptible region IV (L176–R215). The two protease susceptible regions together correspond closely to the segment (S136–R215) separating strands β1 and β3 (Fig. 1a). Because this segment includes IRS3, IRS4, IRS5, and IRS6, we propose local unfolding follows dissociation. In contrast, the C-terminal end of C1 (dark blue), which encompasses strands β3–β5, is stable in the unfolding intermediate. **e:** The major digestion site at R135 is revealed after separation/unfolding of both the N-terminal domain and second β-sheet region (red and magenta in D).

(P200, L201, A204, S205, and G206) situated at the interface between domains (Fig. 3c). Again, there is an overlap between residues within this segment and residues situated at the monomer-monomer interface (IRS6). No unfolding was observed in the final segment (A339, R341, L342, and G343).

The situation of residues from IRS3 (Y140, T141, L142, P144, and H145) at the interface between domains provides an alternative explanation for the extensive disruption of local structure within protease susceptible region III. In contrast, neither disruption of the monomer-monomer interface nor disruption of the interface between domains explain local unfolding around the major digestion site at R135. The surrounding residues are not situated at either the monomer-monomer interface or the interface between domains. Indeed, the side-chain and main-chain atoms of the surrounding residues (V131–I137) are completely buried in a (hypothetical) native subunit from MM-CK and remain buried after separation of the two domains. Nevertheless, in the unfolding intermediate, residues around the major digestion site at R135 (R130–S141) are (relatively) unstructured and exposed at the surface. The affected residues include those at the C-terminal end of strand β1 (V126–R135) of native MM-CK (Fig. 1a). This strand is situated at the center of the large β-sheet that constitutes the core of the C-terminal domain (Fig. 3d). A study of protease substrate binding suggests that, in the absence of gross conformational change, digestion is only possible within the edge strand of an antiparallel β-sheet.[11] Although the digestion results indicate a disruption within β1, they do not support a gross unfolding of the large β-sheet.

The anomalous digestion within β1 may be explained by a topological discontinuity within the large antiparallel β-sheet. Although β1 is first in sequence, this strand bisects the large β-sheet (Fig. 1b). Hence, although not adjacent in sequence to either, β1 (V126–R135) is situated between β5 (L235–K242) and β8 (R292–K298). We propose that strand β1 divides the β-sheet into separate folding units. The results provided by a routine used to identify independent hydrophobic folding units confirm our proposals. The method used evaluates the quality of hydrophobic folding units according to measures of the exposure of nonpolar atoms, compactness, and isolatedness.[20] The results indicate that each CK subunit is composed of three independent hydrophobic folding units (Fig. 1b and Fig. 3d). The first unit (T6–K105) corresponds to the small structural domain at the N-terminus. The two remaining hydrophobic folding units (D119–L127 and G294–K381; S128–G293) each include half of the large β-sheet. Furthermore, β1 (V126–R135) is partitioned (V126–L127; S128–R135) between these two hydrophobic folding units (Fig. 1b). It is of interest that a C-terminal fragment comprising residues Y168–K380 of Mi_b-CK, Y173–K381 of MM-CK, is capable of autonomous folding.[32] Considering the structure of MM-CK, the two halves of the large β-sheet, β2–β5 (G171–K242) and β8–β10 (R292–N338), must indeed be capable of independent folding when the central strand, β1 (V126–R135), is absent.

We conclude that two hydrophobic folding units, or subdomains, together comprise the C-terminal domain (Fig. 3d). In our model, the two subdomains remain loosely associated in the unfolding intermediate. A (partial) separation of the two subdomains would then be required to completely expose β1 at the surface of the unfolding intermediate. However, although separation of the two subdomains is transient under the prevailing conditions, complete separation is induced at higher concentrations of GdmCl. We note that residues at the C-terminal end of β1, including the major digestion site at R135, may be exposed without separation of the subdomains (Fig. 3e). This occurs after separation/unfolding of the N-terminal domain and disruption of the small β-sheet region. However, complete separation is probably necessary for docking with the protease and subsequent cleavage.[11]

The positioning of the remaining digestion sites is consistent with our model of the unfolding intermediate (Fig. 1). Although six major digestion sites (protease susceptible region III; Table I) separate strands β1 (V126–R135) and β2 (G171–P175), two minor digestion sites (protease susceptible region V; Table I) separate strands β5 (L235–K242) and β8 (R292–K298). The region that includes the two (late) minor digestion sites (E262 and E275) comprises a subdomain linker segment likely to promote adhesion. There are six additional minor digestion sites (protease susceptible region IV; Table I) situated in the C-terminal domain, and these separate strands β2 (G171–P175) and β3 (G216–N220). These six additional minor digestion sites suggest that, although strands β2–β5 are included within the same subdomain, strands β3–β5 comprise a more stable unit (Fig. 3d). Because the region separating strands β2 and β3 (protease susceptible region IV; Table I) includes IRS5, IRS6 and IRS7, local structure is likely disrupted during dissociation. In contrast, although the second subdomain does contain many target residues, none were revealed as sites of proteolysis in the unfolding intermediate (with the possible exception of unidentified digestion sites at the extreme C-terminus).[21] The protection of this subdomain from digestion suggests local structure is not disrupted in the unfolding intermediate.

Our proposal that folding (unfolding) of each M-CK subunit proceeds through the adhesion (separation) of autonomous subdomains is consistent with the results of earlier studies. The two tryptophan residues (W218 and W228) situated within the first subdomain (S128–G293) of the C-terminal domain are part of a (nonpolar) denaturation-resistant domain.[5] This region possesses a "relaxed" tertiary structure in the unfolding intermediate.[2,4] There is little evidence in the literature concerning the state of the second subdomain in the unfolding intermediate. This may be partly due to an absence of reporter residues (C, W, and Y) from this region.[4] However, an antibody to a conformational epitope within this subdomain binds both native MM-CK and the unfolding intermediate.[14] In contrast, an antibody against a sequential epitope in the C-terminal "tail" recognizes only denatured BB-CK.[35] Thus, the native conformation of the second subdomain

(G294–K381) is also maintained in the unfolding intermediate. This is consistent with an equilibrium unfolding study of proteinse K-nicked MM-CK. Although native MM-CK is digested with proteinase K at a unique site in a large surface loop between strands β9 and β10, the fragments remain associated below 0.9 M GdmCl.[4,36–38]

Partial Unfolding of Solvent-Exposed Helices

In a native protein, surface loops, helices, and, possibly, the end strands of a β-sheet, may, after local unfolding, adopt the conformation necessary to bind the active site of a protease prior to proteolysis.[11] Our results concerning the unfolding intermediate of MM-CK indicate that local unfolding is largely confined to segments that, in native MM-CK (Fig. 1a), either include helical residues (K32, E37, Y39, E80, R148, R151, R152, E166, E181, E183, L201, and E262) or are situated between elements of secondary structure (F68, Y140, and R209). The principal exception involves the major digestion site (R135) situated within a segment that includes the C-terminal end of β1 (Section (b)-Results and Discussion). In addition, two (late) minor digestion sites (L176 and R215) indicate additional, albeit more limited, unfolding of the β-sheet (Fig. 1a). However, in both instances, the surrounding residues are revealed at the monomer-monomer interface after dissociation (Table II).

Our observation that proteolysis is largely confined to helical segments also agrees with independent experimental data. A large decrease (≈50%) in molar ellipticity at 222 nm is observed on transition from native MM-CK to the unfolding intermediate.[2–4] The reduction in this parameter reflects an extensive loss of native helix. In contrast, because molar ellipticity at 222 nm does not report directly on the β-content of the unfolding intermediate, there is no evidence for a disruption of the β-sheet. A gross disruption of helical structure is consistent with an unfolding of the α-helical N-terminal domain. However, helices in the C-terminal domain are also susceptible to unfolding. The solvent-exposed termini of these helices seem to be particularly affected. This "fraying" is consistent with previous observations.[39] To summarize, we have identified digestion sites situated within, or close to the termini of, 7 (α2, α3, α5, α7, α8, α9, and α10) of 12 native α-helices.

CONCLUSIONS

We have used our proteolysis data to produce a structural model for the principal intermediate populated during the equilibrium unfolding of MM-CK. In our model, the α-helical N-terminal domain is significantly disrupted. The sequential unfolding of the N-terminal domain results in the exposure of a complementary nonpolar surface on the C-terminal domain. This provides a possible explanation for the observed solvent-exposure of nonpolar surface. In contrast, the α/β C-terminal domain maintains a native-like structural core. In particular, there is no evidence for a significant disruption of the large β-sheet. In native MM-CK, the β-sheet is partly shielded from solvent by amphipathic α-helices that pack against each face. In the unfolding intermediate, the termini of some of these α-helices are disrupted. A possible consequence is the exposure of nonpolar atoms of residues within the β-sheet to solvent. This model is reminiscent of the "non-uniform expansion" model of the molten globule.[40] However, our results indicate that the unfolding intermediate comprises two loosely associated subdomains. We conclude that our model most resembles a "partly folded state."[41] This state is characterized by subdomains surrounded by, or connected by, unfolded segments. A "Hierarchical Cooperative" process is likely to govern the folding of proteins composed of subdomains and domains[42] and each MM-CK subunit may fold through such a process. In support of this, the principal equilibrium unfolding intermediate of MM-CK possesses a similar structure to a kinetic intermediate transiently populated during refolding.[14,21,36,43]

Transition to the principal unfolding intermediate is correlated with subunit dissociation. Thus, during equilibrium unfolding of MM-CK, dissociation precedes gross unfolding of the subunits. However, significant disruption of the monomer-monomer interface occurs during dissociation. Indeed, disruption of the monomer-monomer interface accounts for many of the changes observed during transition to the unfolding intermediate. We conclude that intersubunit interactions significantly stabilize the monomer-monomer interface. A network of intermolecular hydrogen bonds (and an ion pair) centered on the hotspot (S147–A153) may be important in this respect.

ACKNOWLEDGMENTS

The authors thank Dr. Simon Caper and Prof. John Kay (Department of Biochemistry, Cardiff University) for their extremely generous help in improving our original RasMol images by using their INSIGHT II system.

REFERENCES

1. Grossman SH. An equilibrium study of the dependence of secondary and tertiary structure of creatine kinase on subunit association. Biochim Biophys Acta 1994;1209:19–23.
2. Couthon F, Clottes E, Ebel C, Vial C. Reversible dissociation and unfolding of dimeric creatine kinase isoenzyme MM in guanidine hydrochloride and urea. Eur J Biochem 1995;234:160–170.
3. Gross M, Lusting A, Wallimann T, Furter R. Multiple-state equilibrium unfolding of guanidino kinase. Biochemistry 1995;34: 10350–10357.
4. Clottes E, Leydier C, Couthon F, Marcillat O, Vial C. Denaturation by guanidinium chloride of dimeric MM-creatine kinase and its proteinase K-nicked form: evidence for a multiple-step process. Biochim Biophys Acta 1997;1338:37–46.
5. Clottes E, Vial C. Discrimination between the four tryptophan residues of MM-creatine kinase on the basis of the effect of N-Bromosuccinimide on activity and spectral properties. Arch Biochem Biophys 1996;329:97–103.
6. Couthon F, Clottes E, Vial C. High salt concentrations induce dissociation of dimeric rabbit muscle creatine kinase: physicochemical characterization of the monomeric species. Biochim Biophys Acta 1997;1339:277–288.
7. Polverino de Laureto P, De Fillipis V, Bello MD, Zambonin M, Fontana A. Probing the molten globule state of α-lactalbumin by limited proteolysis. Biochemistry 1995;34:12596–12604.
8. Polverino de Laureto P, Toma S, Tonon G, Fontana A. Probing the structure of human growth hormone by limited proteolysis. Int J Peptide Protein Res 1995;45:200–208.
9. Fontana A, Zambonin M, Polverino de Laureto P, De Filippis V, Clementi A, Scaramella E. Probing the conformational state of apomyoglobin by limited proteolysis. J Mol Biol 1997;266:223–230.

10. Fontana A, Polverino de Laureto P, De Filippis V, Scarmella E, Zambonin M. Probing the partly folded states of proteins by limited proteolysis. Folding Design 1997;2:17–26.
11. Hubbard SJ, Fisenmenger F, Thornton JM. Modelling studies of the change in conformation required for cleavage of limited proteolytic sites. Protein Sci 1994;3:757–768.
12. Lecomte JTJ, Kao Y-H, Cocco MJ. The native state of apomyoglobin described by proton NMR spectroscopy: the A-B-G-H interface of wild-type sperm whale apomyoglobin. Proteins Struct Funct Genet 1996;25:267–285.
13. Gilmanshin R, Dyer RB, Callender RH. Structural heterogeneity of the various forms of apomyoglobin: implication for protein folding. Protein Sci 1997;6:2134–2142.
14. Morris GE. Monoclonal antibody studies of creatine kinase. The ART epitope: evidence for an intermediate in protein folding. Biochem J 1989;257:461–469.
15. Morris GE, Jackson PJ. Identification by protein microsequencing of a proteinase V8 cleavage site in a folding intermediate of chick muscle creatine kinase. Biochem J 1991;280:809–811.
16. Fritz-Wolf K, Schnyder T, Wallimann T, Kabsch W. Structure of mitochondrial creatine kinase. Nature 1997;381;341–345.
17. Rao JKM, Bujacz G, Wlodawer A. Crystal structure of rabbit muscle creatine kinase. FEBS Lett 1998;439:133–137.
18. Sobolev V, Sorokine A, Prilusky J, Abola EE, Edelman M. Automated analysis of interatomic contacts in proteins. Bioinformatics 1999;15:327–332.
19. Hooft RWW, Sander C, Vriend G. Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. Proteins Struct Funct Genet 1996;26:363–376.
20. Tsai CJ, Nussinov R. Hydrophobic folding units derived from dissimilar monomer structures and their interactions. Protein Sci 1997;6:24–42.
21. Webb T, Jackson PJ, Morris GE. Protease digestion studies of an equilibrium intermediate in the unfolding of creatine kinase. Biochem J 1997;321:83–88.
22. Morris GE, Cartwright AJ. Monoclonal antibody studies suggest a catalytic site at the interface between domains in creatine kinase. Biochim Biophys Acta 1990;1039:318–322.
23. Morris GE, Nguyen thi Man. Changes at the N-terminus of human brain creatine kinase during a transition between inactive folding intermediate and active enzyme. Biochim Biophys Acta 1992;1120:233–238.
24. Muhlebach SM, Gross M, Wirz T, Wallimann T, Perriard JC, Wyss M. IV-2 sequence homology and structure predictions of the creatine kinase isoenzymes. Mol Cell Biochem 1994;133/134:245–262.
25. Jones S, Thornton JT. Protein–protein interactions: a review of protein dimer structures. Prog Biophys Mol Biol 1995;63:31–64.
26. Webb T. An intermediate in the unfolding and refolding of creatine kinase. Ph.D. Thesis (Salford University, UK) 1998.
27. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. J Mol Biol 1998;280:1–9.
28. Morris GE, Frost LC, Newport PA, Hudson N. Monoclonal antibody studies of creatine kinase: antibody-binding sites in the N-terminal region of creatine kinase and effects of antibody on enzyme refolding. Biochem J 1987;248:53–59.
29. Gross M, Furter-Graves EM, Wallimann T, Eppenberger HM, Furter R. The tryptophan residues of mitochondrial creatine kinase: roles of Trp-223, Trp-206, and Trp-264 in active-site and quaternary structure formation. Protein Sci 1994;3:1058–1068.
30. Perraut C, Clottes E, Leydier C, Vial C, Marcillat O. Role of quaternary structure in muscle creatine kinase stability: tryptophan 210 is important for dimer cohesion. Proteins Struct Funct Genet 1998;32:43–51.
31. Min KL, Stephens JP, Henry R, Doutheau A, Collombel C. Identification of the creatine binding domain of creatine kinase by photoaffinity labelling. Biochim Biophys Acta 1998;1387:80–88.
32. Gross M, Wyss M, Furter-Graves EM, Wallimann T, Furter R. Reconstitution of active octameric mitochondrial creatine kinase from two genetically engineered fragments. Protein Sci 1996;5:320–330.
33. Mahowald TA. Identification of an ϵ amino group of lysine and a sulfhydryl group of cysteine near the reactive cysteine residue in rabbit muscle creatine kinase. Fed Proc 1969;28:601.
34. Zhou H, Zhang X-H, Yin Y, Tsou C-L. Conformational changes at the active site of creatine kinase at low concentration of guanidinium chloride. Biochem J 1993;291:103–107.
35. Nguyen thi Man, Cartwright AJ, Andrews KM, Morris GE. Structural changes in the C-terminal region of human brain creatine kinase studied with monoclonal antibodies. Biochim Biophys Acta 1991;1076:245–251.
36. Leydier C, Clottes E, Couthon F, Marcillat O, Ebel C, Vial C. Evidence for kinetic intermedate states during the refolding of GdnHCl-denatured MM-creatine kinase: characterization of a trapped monomeric species. Biochemistry 1998;37:17579–17589.
37. Price NC, Murray S, Milner-White EJ. The effect of limited proteolysis on rabbit muscle creatine kinase. Biochem J 1981;199:239–244.
38. Leydier C, Anderson JS, Couthon F, et al. Proteinase K processing of rabbit muscle creatine kinase. J Protein Chem 1997;16:67–74.
39. Pappenberger G, Schurig H, Jaenicke R. Disruption of an ionic network leads to accelerated thermal denaturation of D glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium Thermatoga Maritima. J Mol Biol 1997;274:676–683.
40. Ptitsyn OB. Molten globule and protein folding. Adv Protein Chem 1995;47:83–217.
41. Freire E. Thermodynamics of partly folded intermediates in proteins. Annu Rev Biochem 1995;24:141–165.
42. Freire E, Murphy KP. Molecular basis of cooperativity in protein folding. J Mol Biol 1991;222:687–698.
43. Grossman SH, Pyle J, Steiner RJ. Kinetic evidence for active monomers during the reassembly of denatured creatine kinase. Biochemistry 1981;20:6122–6128.