

Comparative molecular modeling of *Amphioxus* calcium vector protein with calmodulin and troponin C

Jos A.Cox¹, Philippe Alard² and Olivier Schaad

Department of Biochemistry, University of Geneva, 1211 Geneva 4, Switzerland and ²Unité de Conformation de Macromolécules Biologiques, Université Libre de Bruxelles, CP 160, Avenue Héger, 1050 Brussels, Belgium

¹To whom correspondence should be addressed

Calcium vector protein (CaVP), a new protein isolated from *Amphioxus* muscle, binds in a Ca²⁺-regulated manner to a 27 kd target protein, named CaVPT, whose function has not been elucidated yet. CaVP bears significant sequence homology to both calmodulin and skeletal muscle troponin C, especially in the C-terminal half of the molecule, which presumably contains the two functional Ca²⁺-binding sites. The N-terminal half contains two abortive EF-hands and is intramolecularly crosslinked with a disulfide bond. Using the crystallographic structures of calmodulin and striated muscle troponin C as a framework, we constructed two different three-dimensional models of CaVP and modeled the intramolecular disulfide bridge. The modeling based upon the coordinates of calmodulin yields a Ca²⁺-filled sites configuration in the N-terminal half of the molecule, even though no Ca²⁺ is bound in this half, whereas the troponin C-derived model generates a Ca²⁺-empty sites configuration. The models predict that neither in the Ca²⁺-filled nor in the Ca²⁺-empty sites conformation is there any steric and/or energetic obstacle for the formation of the disulfide bridge and that the disulfide bond is poorly accessible to reducing reagents. The optical properties of the Trp and Tyr residues of CaVP indicate that the calmodulin-derived model represents the most plausible prediction.

Key words: calcium-binding protein/disulfide bridge/ EF-hand/ molecular modeling

Introduction

The muscle of *Amphioxus*, an invertebrate animal with the morphology of a fish, contains large amounts of a unique Ca²⁺-binding protein, abbreviated CaVP (Cox, 1986). The protein binds 2 Ca²⁺ ions with a K_d of 0.11 μ M and interacts in a Ca²⁺-dependent manner with amphiphilic model peptides and with phenyl Sepharose. Qualitatively these interactions are similar to those of other well-known intracellular Ca²⁺ vectors such as calmodulin (CaM) and troponin C (TnC) (Comte *et al.*, 1983). *In vivo* most of the CaVP participates in a 50 kd complex with an endogenous protein (Cox, 1986).

The elucidation of the amino acid sequence of CaVP by Kobayashi *et al.* (1987) revealed four repeats of the so-called EF-hand domain (Kretsinger, 1987). Domains III and IV (defined as in Table I) are canonical according to rules established for >200 EF-hand domains and show high sequence identity with the corresponding domains of CaM and of chicken skeletal muscle TnC. Since CaVP binds only two Ca²⁺, it is presumed that domains III and IV are 'active'. Domain III contains two ϵ -trimethyllysine residues in the α -helices flanking the

Ca²⁺-binding loop. In its native form CaVP contains a disulfide bridge linking the first α -helix of domain I to the second of domain II. The optically interesting aromatic chromophores Trp80, Trp87 and Tyr53 are located in the nonfunctional domain II or in the linker segment between domains I and II are not sensitive to cation binding (Cox, 1986). The numbering of the residues of CaVP throughout the text refers to the one used in Table I, which starts at the N-terminus of troponin C. The subtraction of 7 from this numbering yields the numbering used in Kobayashi *et al.* (1987).

The crystal structure of CaM with four bound Ca²⁺ has been determined at 2.2 Å resolution (Babu *et al.*, 1988); that of TnC, with two Ca²⁺ bound to domains III and IV, has been reported at 2.0 Å resolution (Herzberg and James, 1988; Satyshur *et al.*, 1988). Interestingly, the orientation of the pair of helices surrounding the loops of domain I and II (the Ca²⁺-empty sites) in the crystal structure of TnC is different from that in the Ca²⁺-filled domains III and IV: in I and II the pairs are arranged in an antiparallel fashion and in III and IV the orientation of the pairs is rather perpendicular (as in CaM). It has been proposed that upon Ca²⁺-binding to these domains the interhelical angle diminishes (Herzberg *et al.*, 1986).

As a step toward the elucidation of the function of CaVP and toward the understanding of the principles underlying the activation of enzymes by Ca²⁺-binding proteins, it is of interest to know the three-dimensional structure of CaVP. However, since CaVP rapidly deteriorates in solution, it will be difficult to grow suitable crystals. The high sequence homology of CaVP with CaM and TnC incited us to deduce its three-dimensional structure from template modeling using the CaM and TnC structures as starting models. Such a study also provides an evaluation of the probability of intramolecular disulfide bridging and its importance in the stabilization of the N-terminal half of CaVP. Solution studies on the Trp and Tyr environment provide the possibility of designating the more likely model.

Computational procedures

As a first step in the modeling, the structures of CaM and TnC, whose coordinates were taken from the Brookhaven Data bank (Bernstein *et al.*, 1977), were modified to cope with the insertion of two residues at position 66–67 and one residue at position 127, and with the deletion of one, and four residues respectively between residues 91 and 92. Since the latter deletions occur in the middle of the central α -helix (see Table I), the structure of the residues surrounding the deletion was modeled by fitting the positions of the α carbons to that of an α -helix. The insertions in CaVP occur in loop regions and were modeled using the 'spare parts' approach of Claessens *et al.* (1989). This procedure includes three steps: (i) two anchor regions were defined on both sides of the locus to be modified; (ii) from a BRUGEL database containing 82 highly refined protein structures (Claessens *et al.*, 1989) backbone fragments with the adequate number of residues were selected so as to link the anchors, without modifying their relative coordinates; (iii) after

Table I. Alignment of amino acid sequence of *Amphioxus* CaVP with those of vertebrate CaM and chicken skeletal muscle TnC

Consensus sequence	E n * * n n * * n D * D * D G * I D * * E n * * n n * * n E									
	X Y Z -Y -X -Z									
	<div style="display: flex; justify-content: space-around; margin-bottom: 5px;"> ← M-helix → ← A-helix → ← B-helix → </div>									
	5	10	15	20	25	30	35	40	45	50
Calmodulin			A D Q L T E E Q I A E F K E A F S L F D K D G D G T I T T K E L G T V M R S L G Q N P T							
CaVP			A A P K A R A L G P E E K D E C M K I F D I F D R N A E N I A P V S D T M D M L T K L G Q T Y T							
Troponin C	A S M T D Q Q A E A R A F L S E E M I A E F K A A F D M F D A D G G G D I S T K E L G T V M R M L G Q N P T									
	<div style="display: flex; justify-content: space-around; margin-bottom: 5px;"> ← C-helix → ← D/E-helix → </div>									
	55	60	65	70	75	80	85	90		
Calmodulin	E A E L Q D M I N E V - - D A D G N G T I D F P E F L T M M A R K M K D T D									
CaVP	K R E T E A I M K E A R G P K G D K K N I G P E E W L T L C S K W V R Q D D									
Troponin C	K E E L D A I I E E V - - D E D G S G T I D F E E F L V M M V R Q M K E D A									
	<div style="display: flex; justify-content: space-around; margin-bottom: 5px;"> ← D/E-helix (cont.) → ← F-helix → </div>									
	95	0	5	10	15	20	25			
Calmodulin	- - - S E E E I R E A F R V F D K D G N G Y I S A A E L R H V M T N L G E K' - L									
CaVP	- - - E E E I L R A F K' V F D A N G D G V I D F D E F K F I M Q K' V G E E P L									
Troponin C	K G K S E E E L E D C F R I F D K N A D G F I D I E E L G E I L R A T G E H - V									
	<div style="display: flex; justify-content: space-around; margin-bottom: 5px;"> ← G-helix → ← H-helix → </div>									
	30	35	40	45	50	55	60	65		
Calmodulin	T D E E V D E M I R E A D I D G D G Q V N Y E E F V Q M M T A K									
CaVP	T D A E V E E A M K E A D E D G N G V I D I P E F M D L I K S K K N A L K E S									
Troponin C	T E E D I E D L M K D S D K N N D G R I D F D E F L K M M E G V Q									

Double dots correspond to identical residues, single dots to conserved replacements
 The disposition of the α -helices is indicated (Strynadka and James, 1989).
 K is trimethyl lysine.

comparison with the parent structure, all or part of the best fitting backbone fragment is inserted and the coordinates registered in 'modified' CaM or TnC data files.

The program package BRUGEL, versions 8 and 9 (Delhaise *et al.*, 1984), has been used to display and manipulate the atomic coordination files of CaM and TnC. All the hydrogen atoms were added in standard positions. Gradually the amino acid residues in both proteins were substituted by those of CaVP followed by a conformational energy search about the dihedral angles χ_1 and χ_2 of this side chain in order to obtain a locally minimized conformation. The mutations where the substituting side chain is smaller than the substituted one were performed first. For the oxidized models we determined in the next step the shortest distance between the sulfur atoms of the two Cys residues by continuous rotation around the two torsional angles and imposed the constraint of an internal disulfide bond. In this way four sets of coordinates were obtained for CaVP: two sets derived from CaM and two from TnC, in both cases with (CaVP_CaM_Ox and CaVP_TnC_Ox) and without (CaVP_CaM_Red and CaVP_TnC_Red) the disulfide constraint.

Energy minimizations were performed on all the structures examined in this report. In the CaVP_CaM_Red and CaVP_TnC_Red models the constraint on the positions of the peptide backbone atoms was kept high. The oxidized forms were minimized with low constraint on the N-terminal 40 residues,

to relieve tension resulting from the S-S bridge formation. For the side chains, we initially minimized under constraint with the steepest descent (Levitt and Lifson, 1969) and the conjugate gradient method (Fletcher and Reeves, 1964). Subsequently we minimized without constraints on the side chain atoms, except for those directly involved in the Ca^{2+} -coordination, with the steepest descent method until the root mean square of forces was $< 1.0 \text{ kcal/mol} \cdot \text{\AA}$. A final minimization was carried out using the conjugate gradient method constraint on the main chain until an RMS of $0.1 \text{ kcal/mol} \cdot \text{\AA}$ was achieved. The solvent accessible surface (SAS) is calculated using a spherical probe of the size of water (1.4 \AA). Its area was calculated using a program written by one of us (P.Alard) for the BRUGEL environment.

Results

Sequence homology between CaVP, CaM and TnC

The optimal alignment of the sequences of CaVP, CaM and TnC, confirmed using a program of multiple sequence alignment with hierarchical clustering (Corpet, 1988), is shown in Table I. The CaVP/CaM and CaVP/TnC pair show 49 and 47 identical residues and 28 and 29 conserved substitutions respectively. The sequence homology was highest in the C-terminal half. ~42% sequence identity and 20% conservative replacements with respect to the

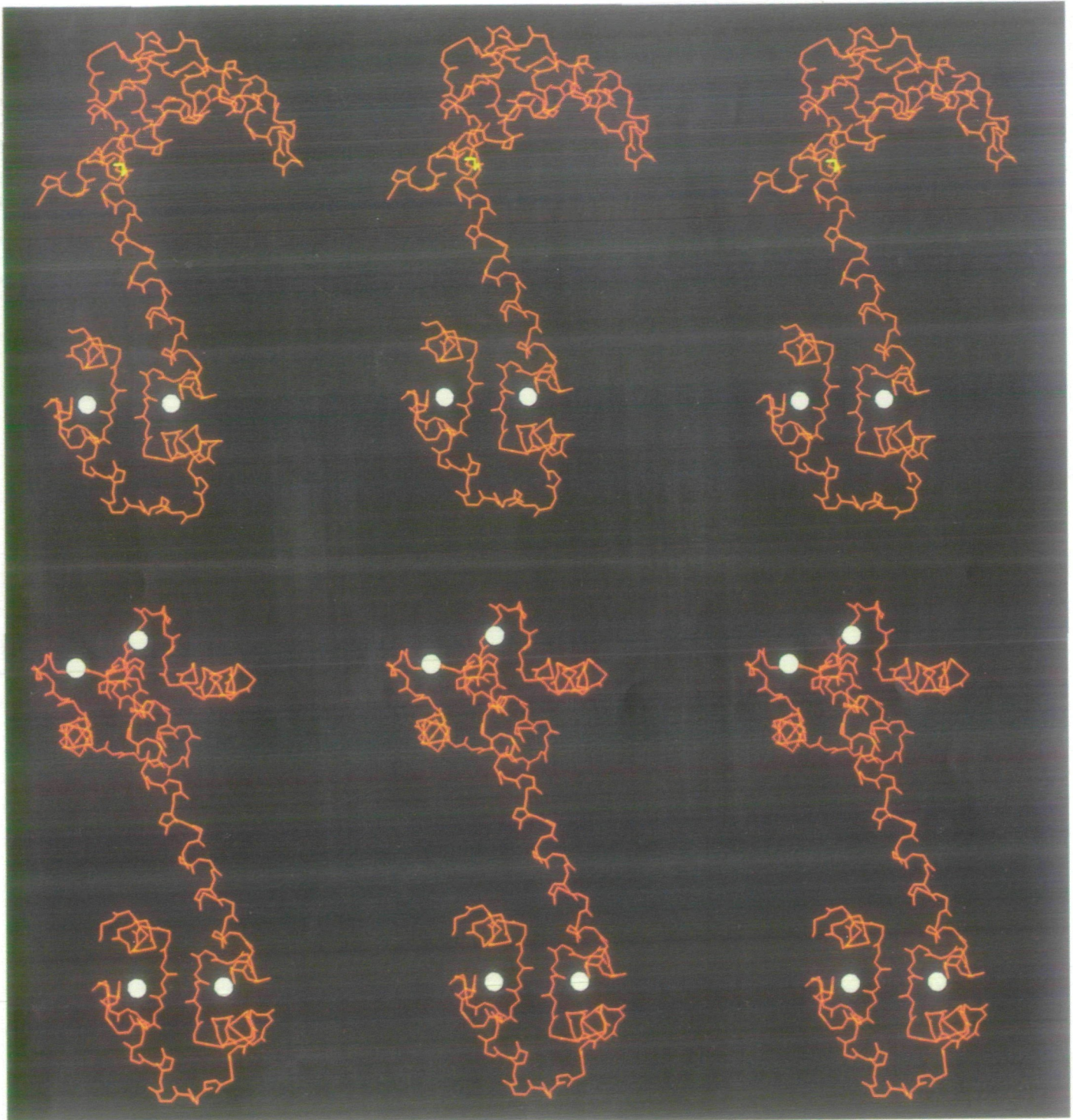


Fig. 1. Stereoscopic views of the peptide backbones of CaVP_CaM_Ox (a) and CaM (b) generated with the Brugel environment on the Evans and Sutherland graphic display. The structures were displayed from the same vantage point, i.e. with the central α -helix and the two C-terminal Ca^{2+} ions in the plane and the H-helix sticking out toward the observer. S-S bridge depicted in yellow. The ionic radius of Ca^{2+} is 1.06 Å.

C-terminal halves of both proteins. The residue numbering indicated in Table I, which starts with the N-terminus of TnC, but subsequently adopts that of CaVP, will be used throughout. The α -helices are designated A, B, C, D/E (= central helix), F, G and H from the N- to the C-terminus (see Figure 1). The additional N-terminal α -helix in the TnC-derived model is designated N-helix.

CaM-derived form of CaVP

Comparison with the parent molecule. Figure 1(a) is the stereopresentation of the α -carbon model of CaVP (segment 10–152) in the oxidized form as seen in front of the C-terminal cup with the two bound Ca^{2+} behind the cup, equidistant to the observer. The original CaM structure, seen from the same vantage point, is displayed in Figure 1(b). The general features are similar

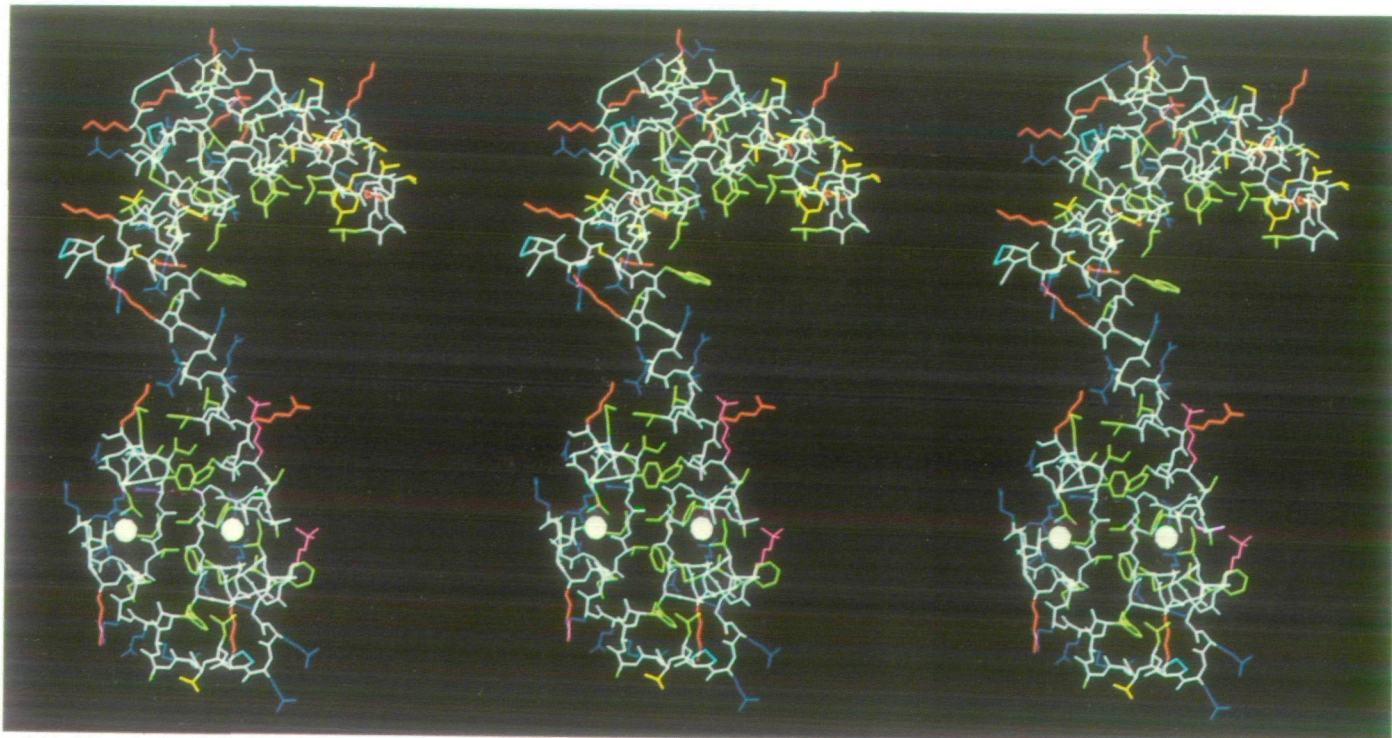


Fig. 2. Stereoscopic view of the complete structure of CaVP_CaM_Ox. The color code is as follows: white, the peptide backbone; red, Asp and Glu; blue, Arg, Lys and His; purple, TML, green, Met, Ala, Val, Leu, Ile, Phe and Trp; orange, Ser, Thr, Tyr, yellow, Cys; light blue, Pro and Gly. N-terminal at top. Trp87 is positioned at the top of the central helix, pointing to the right.

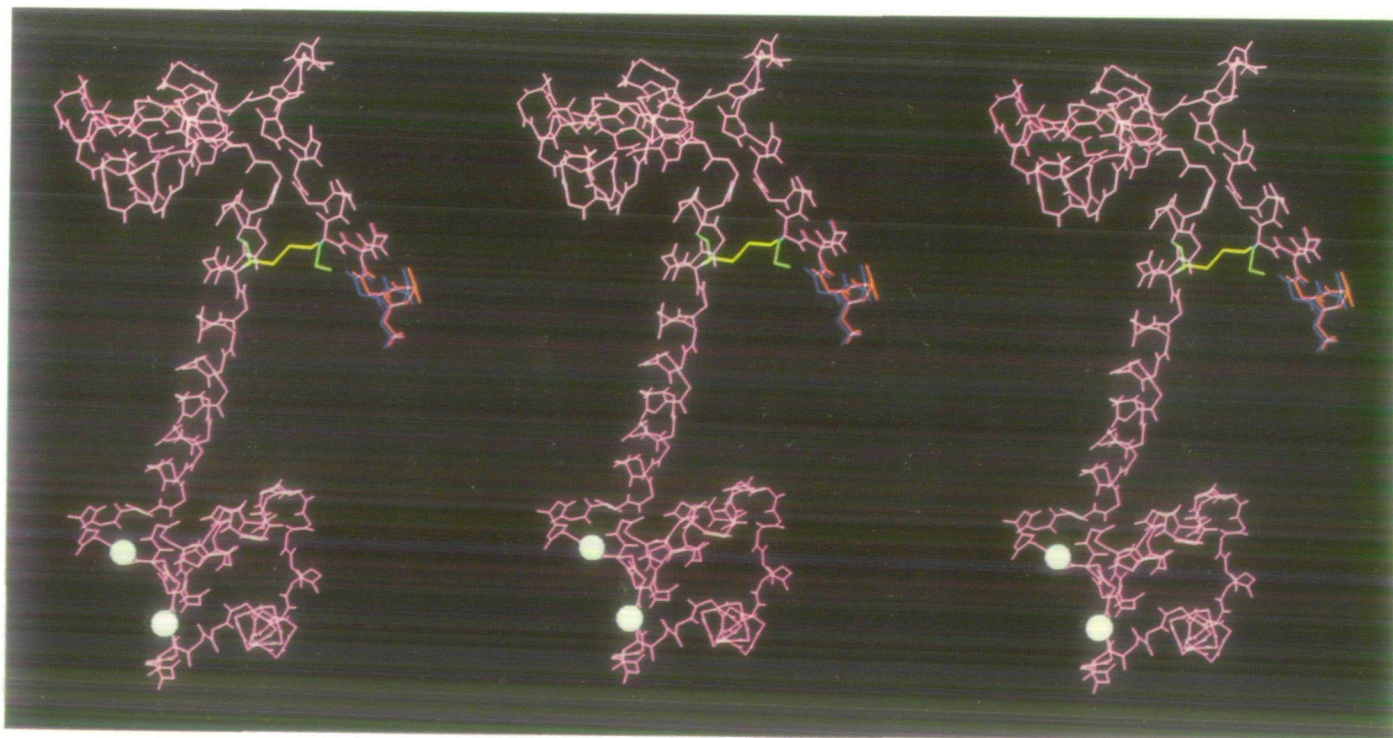


Fig. 3. Stereoscopic and superposed views of CaVP_CaM_Ox (red) and CaVP_CaM_Red (blue). Perfect superposition of the two structures yields purple. Polar hydrogens are displayed. S-S bridge depicted in orange; cysteine side chains in green.

including the dumbbell shape and the presence of the cups in the N- and C-terminal halves. The total length is 61 Å and that of the central α -helix (Glu78 to Phe103) is 38.5 Å. One notices an important difference: the orientation of the N-terminal domain

relative to the fixed C-domain is 80° clockwise as compared to 180° for CaM, thus changing the general form of the protein. The Pro insertion in the loop linking EF-domains III and IV is also clearly visible (lower right in Figure 1a and b). The

Table II. Hydrogen bonds and salt bridges expected to be of particular importance in the ternary structure of CaVP

Type of bridge	CaM__Ox	CaM__Red	TnC__Ox	TnC__Red
Salt	Asp20–Lys24	Asp20–Lys24	Lys19–Asp20	Lys19–Asp20
Salt	Lys69–Glu78	Lys69–Glu78	Lys55–Glu59	Lys55–Glu59
Salt	–	–	Glu57–Lys86	Glu57–Lys86
Twin salt	–	–	Glu93–Lys158	Glu93–Lys158
Salt	Glu134–Lys138	Glu134–Lys138	Lys117–Glu130	Lys117–Asp130
Salt	–	–	Asp155–Lys160	–
Twin Hbo	Ala37–Ile75	Ala37–Ile75	Ala37–Ile75	Ala37–Ile75
Twin Hbo	Ile111–Ile148	Ile111–Ile148	Ile111–Ile148	Ile111–Ile148
Hbo	Arg31–Asp41	Arg31–Asp41	Tyr53–Lys86	Tyr53–Lys86
Hbo	Asn32–Asp41	Asn32–Asp41	–	–
Hbo	Arg66–Lys72	Arg66–Lys72	Arg66–Lys72	Arg66–Lys72
Hbo	Asn106–Glu115	Asn106–Glu115	Ala105–Glu115	Ala105–Glu115
Hbo	Asp143–Glu152	Asp143–Glu152	Asp143–Glu152	Asp143–Glu152

Arg–Gly insertion between EF-hands I and II adds an irregular turn to the C-helix but does not significantly change the shape of the N-terminal. It should be remembered that the nine-residue-long N-terminal and nine-residue-long C-terminal segments are not presented here, since the coordinates are not present in the parent CaM structure. The complete tertiary structure of CaVP derived from CaM is shown in Figure 2. Stabilizing non-covalent bonds include three salt bridges and seven long-range hydrogen bonds listed in Table II. Among them are the characteristic two twin hydrogen bonds in each half, also present in the parent CaM molecule.

The Cys residues are implanted on the sides of α -helices A and D and point to each other (Figure 3). In the reduced form of CaVP the shortest distance between the sulfur atoms is 3.48 Å, which makes disulfide bridge formation possible with minimal distortion of the backbone structure. The largest displacement is indeed the discrete movement of the end of the A-helix. Among the side chains only Pro16 undergoes an important movement (not shown). Table II shows that the stabilizing long-range forces are identical in the oxidized and reduced structures of CaVP and Table III shows that the hydrophobic cup is made up of the same residues with nearly identical SAS values, with the exception of Trp80.

N-terminal half. The N-terminal half can be presented as a ladle with the handle inserted upon the rim, perpendicular to the surface of the cup (Figure 4). The cup contains a water-exposed hydrophobic interior made up of the residues listed in Table III. In the parent CaM molecule also side chains at homologous positions form the hydrophobic cup. The size of the continuous solvent-accessible hydrophobic area within the cup (THS) is very similar in CaVP__CaM__Ox and in CaM. The composition of this area, however, is strikingly different: in CaVP it is composed mostly (>85%) of branched and/or bulky side chains; in CaM >40% consists of the unbranched Met. Met confers considerable flexibility, which has been invoked to explain the multi-functional character of CaM (O'Neil and DeGrado, 1990). Trp87 forms a fully solvent-accessible hydrophobic bud (SAS = 166 Å²) on the central α -helix at the entrance of the cup. The hydrophilic rim is displayed in Figure 5, together with the particular orientation of Tyr53 and of the disulfide bridge. Tyr53 is located at the rim of the cup, opposite to the central α -helix and is mostly solvent-exposed (SAS = ~60 Å²). The S–S bridge lies on the rim at the basis of the central α -helix, opposite to the Tyr-side. The S–S bond is mostly solvent-hidden (SAS = 14 Å²) with access from the inside of the cup (yellow dots at the right in Figure 5). Figure 5 also shows that in contrast to the case of CaM (Kret-

Table III. The composition of the solvent-exposed hydrophobic area inside the cups in CaVP and the contribution of the individual residues.

CaVP__TnC__Ox		Parent TnC	
Ala99	25	Leu98	19
Val102	111	Cys101	29
Phe103	59	Ile104	126
Phe116	23	Phe105	54
Ile119	33	–	–
Met120	57	Ile121	32
Val123	130	Leu122	41
Leu128	37	Val129	35
Ala136	26	Leu137	73
Leu156	65	Met157	50
Ile157	82	Met158	88
THS/cup	648	–	547
TS	10042	–	10001
CaVP__CaM__Ox		Parent CaM	
–	–	Phe12	52
Ile25	76	Ala15	44
Ile28	95	Leu18	101
Phe29	44	Phe19	45
Met45	24	Leu32	21
Leu46	31	Met36	39
Leu49	129	Leu39	135
Ala60	62	Met51	57
Ile61	64	Met71	43
Trp80	57	Met72	43
Leu83	65	Met76	96
THS/cup	647	–	676
Ala99	36	Ile85	33
Val102	87	Ala88	36
Phe103	38	Val91	106
Phe116	47	Phe92	29
Ile119	22	Leu105	16
Met120	57	Met109	39
Val123	137	Leu112	99
Leu128	30	Leu116	27
Ala136	35	Met124	80
Ala140	17	–	–
Leu156	76	Met144	85
Ile157	89	Met145	89
THS/cup	671	–	639
TS	10147	–	11282

TS, total surface of the protein in Å²; THS/cup, total accessible hydrophobic surface per cup.

For the parent CaM and TnC the numbering is that of the original papers.

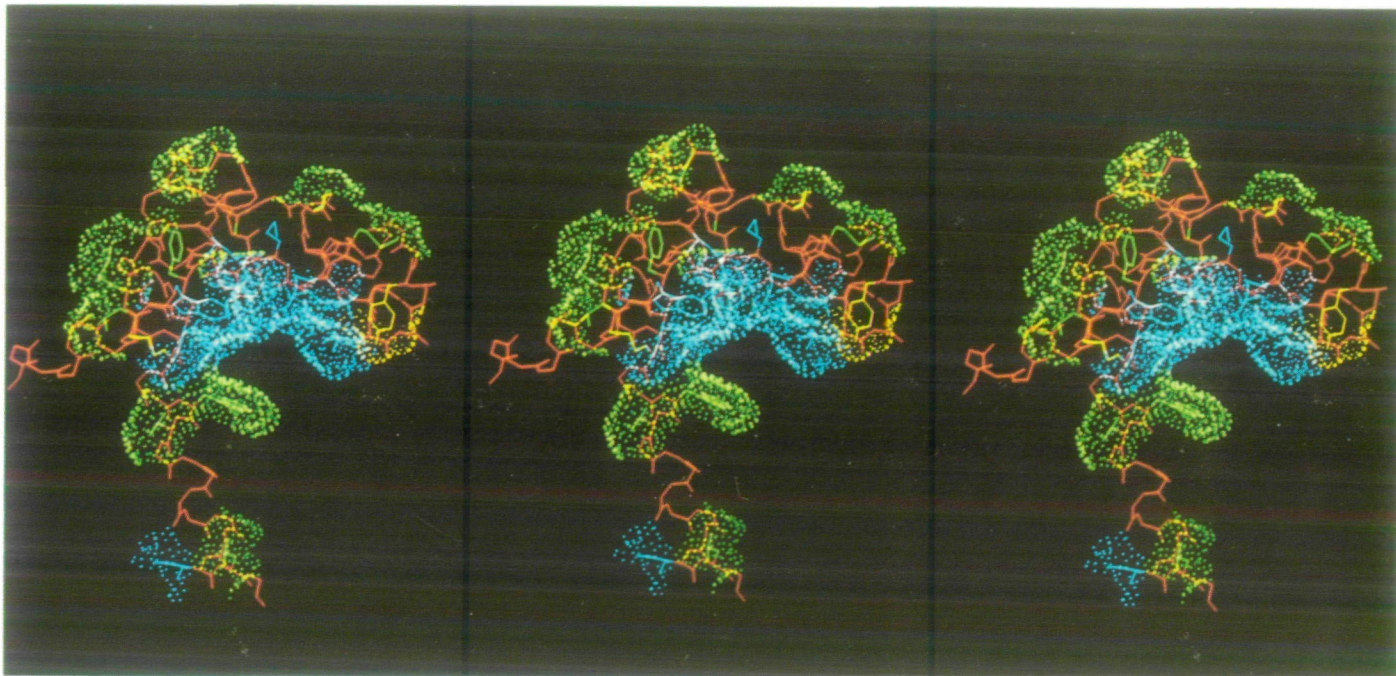


Fig. 4. Stereoscopic view of the N-terminal half of CaVP_CaM_Ox (segment 15-98) with display of all hydrophobic side chains and their solvent-accessible areas. Those forming isolated patches are in green; those forming the continuous hydrophobic area in the cup are in light blue. S-S bridge depicted in yellow; Tyr53 in orange at the right.

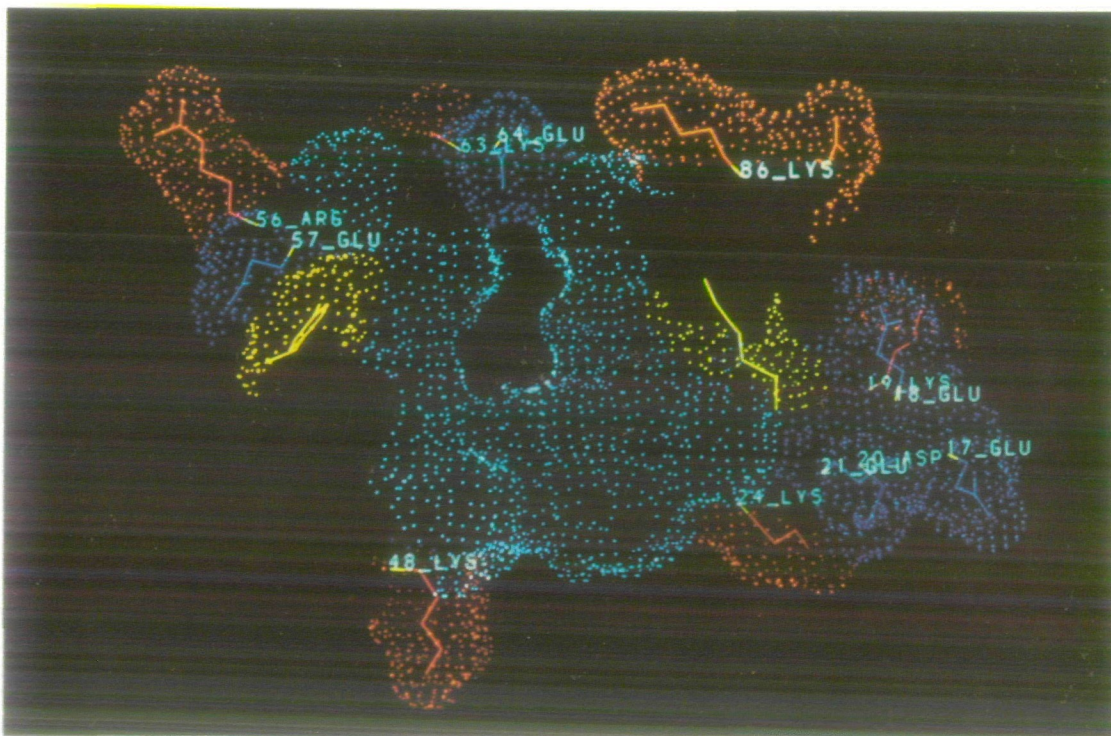


Fig. 5. View of the side chains of particular residues at the rim of the N-terminal cup of CaVP_CaM_Ox together with their accessible surface. The picture is taken in front of the cup with the continuous hydrophobic area (see Figure 5) in light blue and the central α -helix, implanted at the lower right side, directed towards the observer. At the rim are displayed the basic (red) and acidic (blue) side chains, as well as Tyr53 (orange) and the S-S bridge (yellow). Only side chains are displayed; small yellow sticks mark their insertion onto the α -carbons.

singer *et al.*, 1986; Weber *et al.*, 1989), the charge balance on the rim is rather neutral.

C-terminal half. The C-terminal half (Figure 6) is very similar

to that of CaM and can be presented as a spoon with the handle inserted upon the rim with an angle of 145° with respect to the plane of the rim. The two functional Ca^{2+} -binding loops sit on

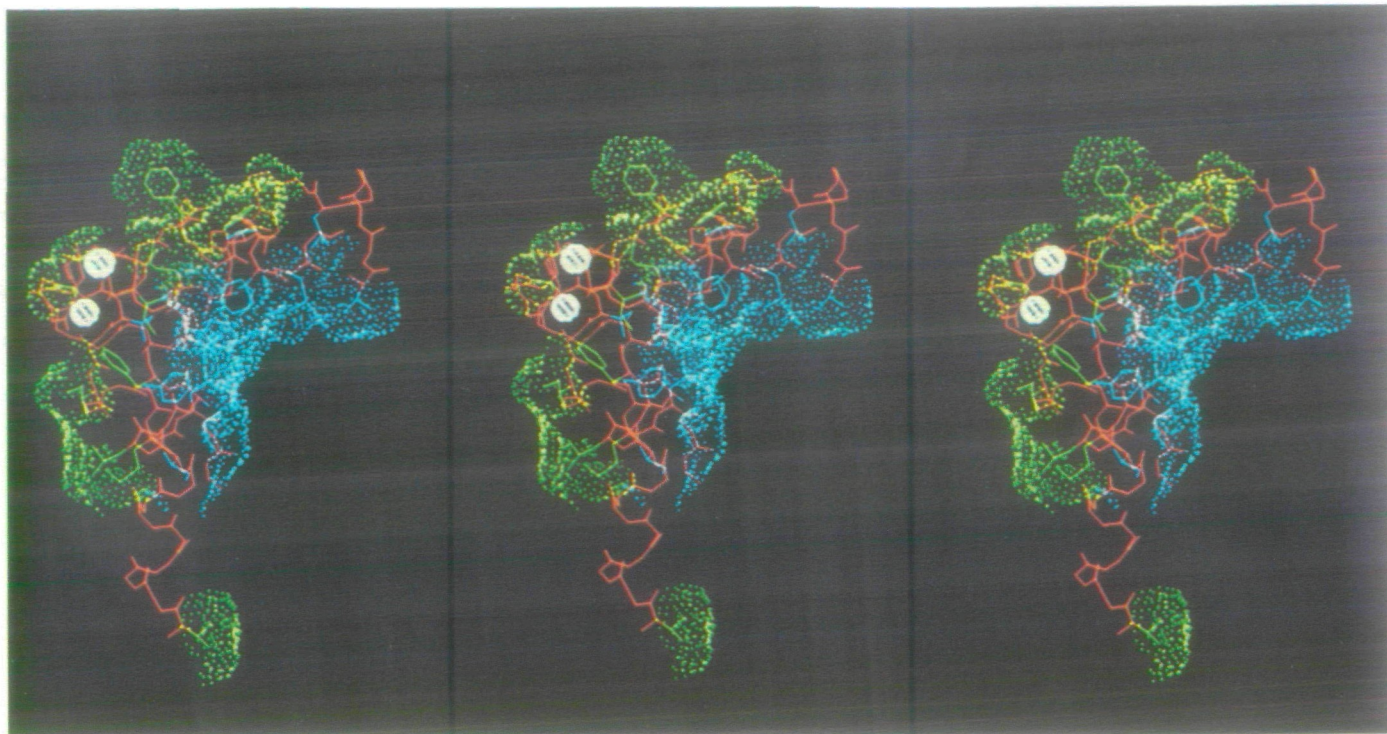


Fig. 6. Stereoscopic view of the C-terminal half of CaVP_CaM_Ox (segment 88–158) with display of all hydrophobic side chains and their solvent-accessible areas. Those forming isolated patches are in green; those forming the continuous hydrophobic area in the cup are in light blue. Two Ca^{2+} -sites are located at the back of the spoon-like structure.

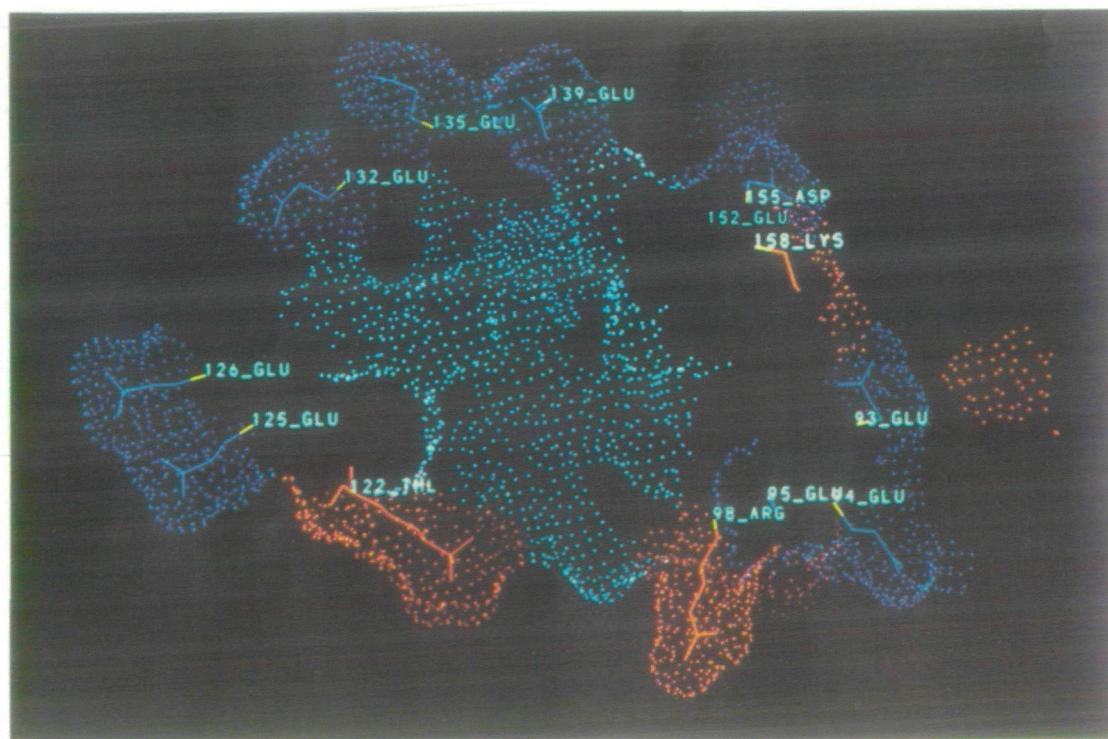


Fig. 7. View of the side chains of particular residues at the rim of the C-terminal cup of CaVP_CaM_Ox together with their accessible surface. The picture is taken in front of the cup with the continuous hydrophobic area in light blue and the central α -helix, implanted at the lower right, slightly bent upwards to the observer (spoon configuration). At the rim are displayed the basic (red) and acidic (blue) side chains. Only side chains are displayed; small yellow sticks mark their insertion onto the α -carbons.

the backside of the cup and their short antiparallel β -pleated sheet segments are stabilized by twin hydrogen bonds between the main chain NH and CO groups of Ile111 and Ile148. The

Ca^{2+} – Ca^{2+} distance is 11 Å and the average Ca^{2+} to oxygen ligand distance is 2.2 Å. The solvent-accessible hydrophobic surface inside the cup is similar in size to that of CaM, but again

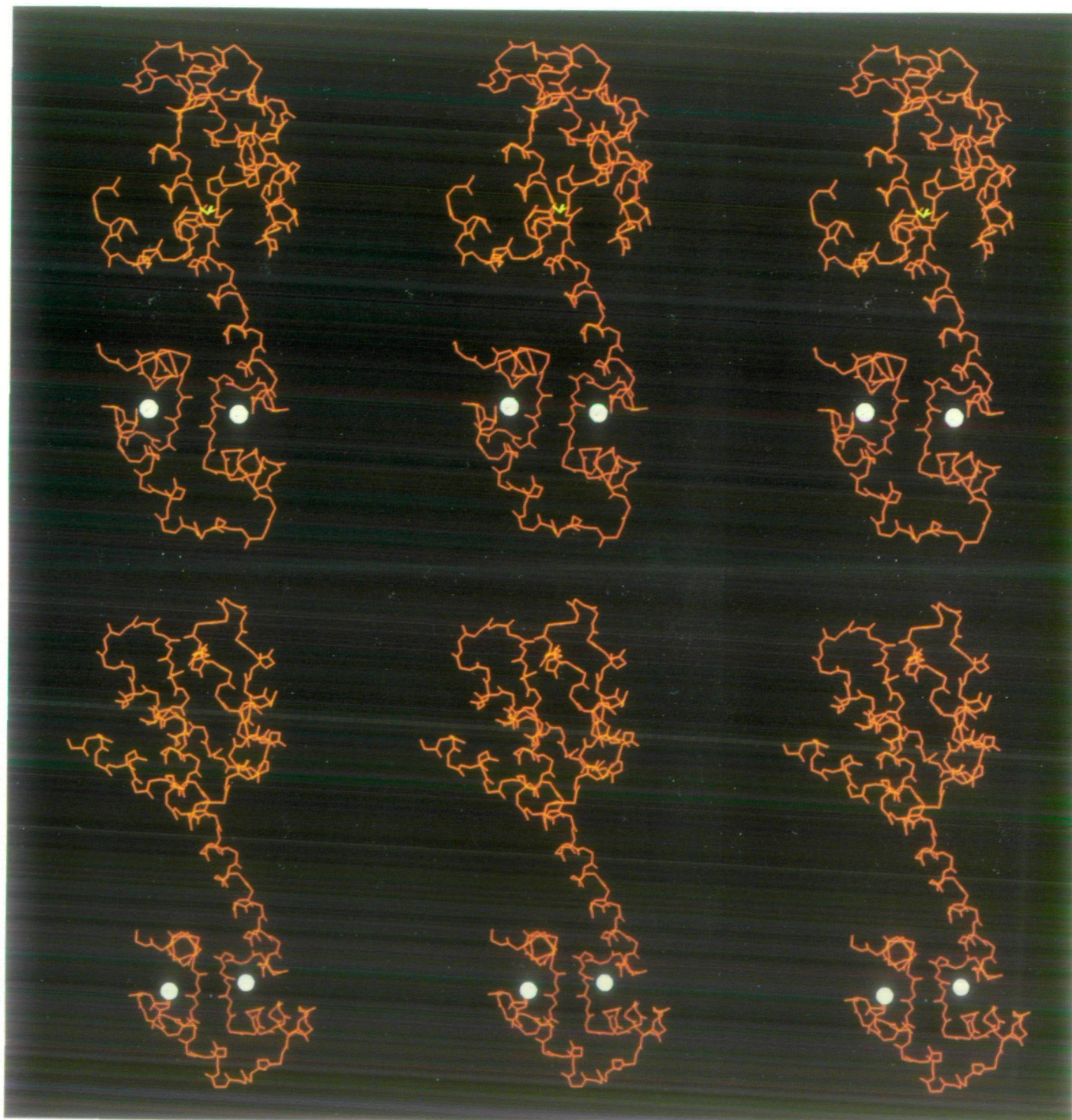


Fig. 8. Stereoscopic views of the peptide backbones of CaVP__TnC__Ox (a) and TnC (b). The structures were displayed from the same vantage point, similar to the one for CaM in Figure 2. The S-S bridge is depicted in yellow.

predominantly composed of branched and/or bulky side chains (Table III). The rim is decorated with a majority of negatively charged residues: 10 negative against four positive residues including Tml122 (Figure 7). The second Tml residue, implanted on the central α -helix, points away from the cup (Figure 2).

TnC-derived form of CaVP

Comparison with the parent molecule. Figure 8(a) is the stereo-presentation of the α -carbon model of CaVP (segment 1 – 154)

in the oxidized form. The original TnC structure is displayed in Figure 8(b). The general features are similar (total length 58 Å, length of the central α -helix 37.7 Å) except the orientation of the N-terminal domain relative to the fixed C-domain which is at 80° clockwise (120° for TnC). As in the CaM-derived model the Pro insertion yields an oval-shaped C-terminal cup. The seven-residue-long C-terminal segment of CaVP is not depicted here. Given the orientation of its origin (Figure 8a) it probably points away from the rest of the protein. Particular stabilizing bonds are listed in Table II. Noticeably, the more compact and

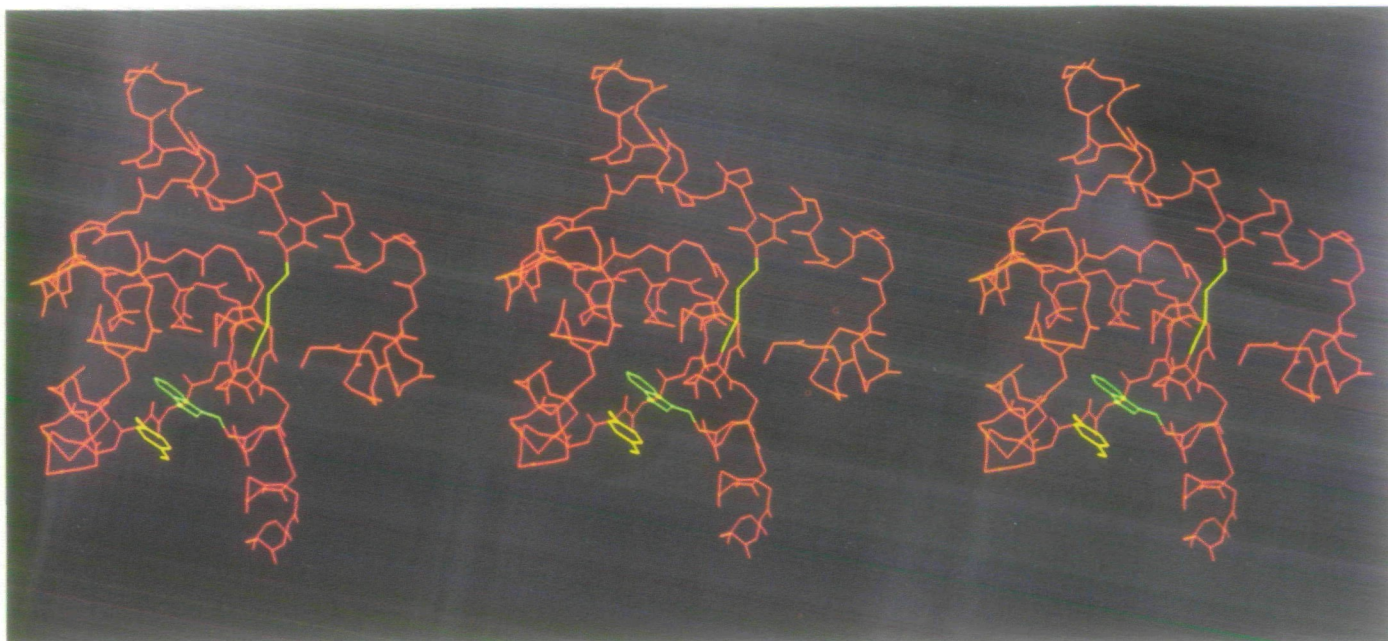


Fig. 9. Stereoscopic view of the N-terminal half of CaVP_TnC_Ox with emphasis of the antiparallel π - π stack between Trp87 (green) and Tyr53 (orange) and the S-S bridge (yellow). Polar hydrogens are displayed.

globular structure of the N-terminal half allows two to three more salt bridges in the TnC-derived model than in that originating from CaM, but the Asn32-Asp41 hydrogen bridge is lost. In the reduced form of CaVP the Cys residues point to each other with a minimal S-S distance of 4.17 Å, which makes disulfide formation possible with little distortion of the structure (data deposited).

N-terminal half. The N-terminal half of the TnC-derived model, in contrast to that derived from CaM, does not possess a distinct solvent-exposed hydrophobic cup: the α -helices are lined up more or less parallel to the central α -helix and there is an additional three-turn α -helix at the N-terminal end, which crosses in front of the central helix. Its overall structure can be compared to that of a mushroom with its cap covering the upper 40% of the central α -helix. Most of the hydrophobic side chains of these amphiphilic α -helices are oriented toward a closed interior. There are two small solvent-exposed hydrophobic patches: Phe13 + Leu13 + Val88 and Ala60 + Ile61 + Leu83, with areas of ~ 125 Å² (not shown). Thus the N-terminal half of TnC-modeled CaVP behaves as a typical polar protein with a hydrophobic core mostly covered with a hydrophobic shield. Within this core and completely shielded from the aqueous solvent are located the three aromatic side chains Tyr53, Trp80 and Trp87. Trp87 and Tyr53 form a π - π stack with antiparallel (177°) oriented aromatic planes (Figure 9). The distance between the geometric centers of the two aromatic groups is 3.3 Å (> 20 Å in the CaM-derived model). The S-S bridge runs through the center of the hydrophobic interior and is also completely shielded from the solvent.

C-terminal half. The C-terminal cup with its hydrophobic interior is well preserved and nearly identical with that of the CaM-derived model. The cup, with its hydrophobic interior and negatively charged rim is very similar to that of the CaM-derived model (data deposited). The continuous hydrophobic surface is distinctly higher than the corresponding surface in the parent TnC.

Discussion

CaVP (18 kd) migrates in Sephadex gel filtration experiments as a 28 kd protein, thus pointing as in the case of CaM and TnC (Cox *et al.*, 1980) to a highly asymmetric shape. Moreover the segment between the second and third Ca²⁺-binding loop shows 47% sequence identity with the corresponding segment in CaM and does not contain α -helix-breaking residues. The predicted 38 Å long central α -helix in CaVP and its dumbbell shape are thus likely. It appears that in both the CaM- and the TnC-derived model, although they represent fundamentally different configurations for the N-terminal half of the molecule, the Cys sulfur atoms are at a distance ideally suited for linkage. The only significant conformational change in both models is a movement by a few ångströms of the N-terminal end towards the central α -helix. Both models readily explain why the disulfide bridge is very resistant to reduction as long as the protein is not denatured and similarly why the disulfide bridge is not easily reformed in an oxidizing environment (Kobayashi *et al.*, 1987). Indeed both the models of the oxidized and reduced forms predict a very poor accessibility of the disulfide bonds or thiol groups to polar reducing or oxidizing reagents. Their accessible surface for a molecule as small as water is minimal (i.e. < 10 Å² and situated in the middle of the hydrophobic cup).

The final and most important question then arises: which model does the real three-dimensional structure of CaVP resemble most: that derived from CaM or that of TnC? As mentioned above, non-discriminating criteria are the inertness of the disulfide bridge and the asymmetric shape. More subtle indications are the optical properties of the Trp and Tyr residues. In the TnC-derived model these three aromatic residues are completely shielded from surrounding water. Moreover Trp87 and Tyr53 form a close π - π stack. In the CaM-derived model, Trp87 and Tyr53 do not form such a stack, since they are far away from each other and completely exposed to water. Trp80 is partially exposed to the solvent. Thus one would expect that in fluorescence and in

near-UV circular dichroic experiments the TnC model should show intense bands with maximal fluorescence closer to 300 than to 350 nm. The CaM model predicts much more solvent quenching, thus small signals and a fluorescence maximum close to 350 nm. The preliminary experimental data available at this time favor the CaM-derived model, since the Tyr and Trp peaks in circular dichroism are small and the protein fluoresces maximally at 337–340 nm (J.A.Cox, unpublished observations). Another discriminating property is the extent of total solvent-accessible hydrophobic surface in both models. The CaM-derived model contains two cups, the TnC-derived model only one. With fluorescent hydrophobic indicators such as 2-(*p*-toluidinyl)naphthalene-6-sulfonate (LaPorte *et al.*, 1980) it becomes possible to determine whether the N-terminal half of CaVP possesses a hydrophobic cleft as predicted by the CaVP_CaM model. It should be noted that, due to the protruding Trp87, the N-terminal hydrophobic cleft in CaVP is much more closed than the N-terminal cleft in the parent CaM molecule. This can explain why the complex with melittin is also less stable (Cox, 1986). Finally, fragment studies will also better assign the model: our preliminary studies (J.A.Cox, unpublished data) indicate fast trypsin proteolysis at Lys86, which is indeed much more solvent exposed in the CaM model than in that derived from TnC. Although these data, fluorescence and circular dichroism favor the CaM model, more detailed experimental information will be required to know unambiguously which model is correct.

Acknowledgement

This work was supported by the Swiss National Science Foundation Grants 31-9162.87 and 3 541-0.86.

References

- Babu, Y.S., Bugg, C.E. and Cook, W.J. (1988) *J. Mol. Biol.*, **204**, 191–204.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) *J. Mol. Biol.*, **122**, 535–542.
- Claessens, M., Van Cutsem, E., Lasters, I. and Wodak, S. (1989) *Protein Engng.*, **2**, 335–345.
- Comte, M., Maulet, Y. and Cox, J.A. (1983) *Biochem. J.*, **209**, 269–272.
- Corpet, F. (1988) *Nucleic Acids Res.*, **16**, 10881–10890.
- Cox, J.A. (1986) *J. Biol. Chem.*, **261**, 13173–13178.
- Cox, J.A., Comte, M. and Stein, E.A. (1980) *Biochem. J.*, **185**, 205–211.
- Delhaise, P., Bardiaux, M. and Wodak, S. (1984) *J. Mol. Graphics*, **2**, 103–106.
- Fletcher, R. and Reeves, C.M. (1964) *Comput. J.*, **7**, 149.
- Herzberg, O., Moulton, J. and James, M.N.G. (1986) *J. Biol. Chem.*, **25**, 2638–2644.
- Herzberg, O. and James, M.N.G. (1988) *J. Mol. Biol.*, **203**, 761–779.
- Kobayashi, T., Takagi, T., Konishi, K. and Cox, J.A. (1987) *J. Biol. Chem.*, **262**, 2613–2623.
- Kretsinger, R.H. (1987) *Cold Spring Harbor Symp. Quant. Biol.*, **70**, 499–510.
- Kretsinger, R.H., Rudnick, S.E. and Weissman, L.J. (1986) *J. Inorg. Biochem.*, **28**, 289–302.
- LaPorte, D.C., Wierman, B.M. and Storm, D.R. (1980) *Biochemistry*, **19**, 3814–3819.
- Levitt, M. and Lifson, S. (1969) *J. Mol. Biol.*, **46**, 269–279.
- O'Neil, K.T. and DeGrado, W.F. (1990) *Trends Biochem. Sci.*, **15**, 59–64.
- Satyshur, K.A., Rao, S.T., Pyzalska, D., Drendel, W., Greaser, M. and Sundaralingam, M. (1988) *J. Biol. Chem.*, **263**, 1628–1647.
- Strynadka, N.C.J. and James, M.N.G. (1989) *Annu. Rev. Biochem.*, **58**, 951–998.
- Weber, P.C., Lukas, T.J., Craig, T.A., Wilson, E., King, M.M., Kwiatowski, A.P. and Watterson, D.M. (1989) *Proteins*, **6**, 70–85.

Received on March 6, 1990; accepted on June 20, 1990