



A bioinformatics study concerning the structural and functional properties of human caveolin proteins

ADRIANA ISVORAN^{1,2*}, DANA CRACIUN³, ALECU CIORSAC⁴, NAHUEL PERROT⁵,
VERONICA BESWICK^{5,6}, PIERRE NEDELLEC⁵, ALAIN SANSON⁵
and NADEGE JAMIN⁵

¹Department of Biology–Chemistry, West University of Timisoara, 16 Pestalozzi, 300316 Timisoara, Romania, ²Laboratory of Advanced Researches in Environmental Protection, West University of Timisoara, 4 Oituz, 300086 Timisoara, Romania, ³Teacher training Department, West University of Timisoara, 4 V. Pirvan, 300223 Timisoara, Romania, ⁴Department of Physical Education and Sport, Politehnica University of Timisoara, 2 P-ta Victoriei, 300306 Timisoara, Romania, ⁵Commissariat à l’Energie Atomique (CEA), Institute of Biology and Technologies of Saclay (iBiTecS) 91191 Gif-sur-Yvette Cedex, France and ⁶Department of Physics, Université d’Evry-val-d’Essonne, 91025 Evry Cedex, France

(Received 16 July, revised 21 September 2013)

Abstract: A bioinformatics study was performed to predict and compare the structural and functional properties of human caveolins: caveolin-1, -2 and -3. The computed local physicochemical properties, predictions of their secondary structure elements and interacting partners of caveolin-2 and -3 were compared to the experimentally determined structural and functional properties of caveolin-1. These data combined with sequence alignments of the three caveolins allowed the functional domains of caveolin-2 and -3 to be predicted and characterised. The hydrophobic regions of these proteins are highly similar in sequences and physicochemical properties, which is in good agreement with their known membrane locations and functions. The most divergent in terms of sequences and properties are the C-terminal regions of the caveolins, suggesting that they might be responsible for their distinct predicted interactions, with direct consequences on signalling processes.

Keywords: secondary structure; disordered regions; functional domains.

INTRODUCTION

Caveolins belong to a family of small (around 20 kDa) integral membrane proteins with both N- and C-termini facing the cytoplasm.¹ This family comprises three members in vertebrates: caveolin-1 (cav1), caveolin-2 (cav2) and caveolin-3 (cav3). Cav1 and cav2 are ubiquitously co-expressed, while cav3 is muscle specific. Caveolins play essential structural roles in the organisation of

*Corresponding author. E-mail: adriana.isvoran@cbg.uvt.ro
doi: 10.2298/JSC130716100I

the caveolae and also participate in many important cellular processes, such as vesicular transport, signal transduction, cholesterol homeostasis and tumour suppression.²

Cav1 is the most studied member of the caveolin family. It is oligomeric and is a cholesterol and fatty acid binding protein. There are two isoforms of cav1: the α - and the β -isoform. The β -isoform lacks the first 32 amino acid residues of the α -isoform.^{3,4} Two types of post-translational modifications have been shown to affect cav1, *i.e.*, palmitoylation and phosphorylation, both of which are important for its functions.^{5,6} Three palmitoylation sites at cysteine 133, 143 and 156 and two phosphorylation sites at tyrosine 14 and serine 80 have been identified.

Caveolin-1 is expressed in most types of cells, whereby adipocytes, endothelial cells, fibroblasts and type I pneumocytes have the highest levels of its expression.⁷ It is found only in vertebrates and has been the subject of numerous research studies, the data of which provide the basis of the present knowledge regarding the interactions of these proteins with their cellular environment. A structural bioinformatics analysis of cav1 revealed its highly conserved sequence among vertebrate organisms (more than 64 % for vertebrates and 99 % for mammals), its high similarity with cav2 and cav3 and enabled to its structural organization and functional regions to be predicted.⁸ The functional regions of Cav1 are the hydrophobic domain (HD), composed of residues 102 to 134, and two adjacent regions flanking the HD: the N-terminal attachment domain (N-MAD, residues 82–101) and the C-terminal attachment domain (C-MAD, residues 135–150), which were found to bind to membrane with high affinity.^{7,9} The C-terminal domain of cav1 contains three cysteine residues that are modified by palmitoylation and are supposed to be involved in caveolin–caveolin and caveolin–lipids interactions.⁷ The N-MAD domain, also known as the scaffolding domain (CSD) of cav1, has been shown to act both as an anchor for different proteins within caveolae and as a regulatory element able to activate or inhibit the signalling activity of a given protein.^{10,11} Moreover, the interactions of cav1 with other proteins may also be mediated by other regions.¹² The CSD includes a short amino acid sequence, residues 94–101, known as the CRAC (cholesterol recognition/interaction amino acid consensus) motif,¹³ which is involved in the membrane interaction of cav1.¹⁴ It was proposed that the CSD has two functional regions: the CRAC region responsible for interactions with cholesterol and its first segment (residues 82–93) being involved in a signal transduction process.⁸

Earlier studies considered that proteins interacting with caveolin have a specific “caveolin binding domain” (CBD), $\psi X \psi \text{XXX} \psi$ or $\psi \text{XXXX} \psi \text{XX} \psi$, where ψ is an aromatic residue (Trp, Tyr or Phe)¹⁵ that interacts with the CSD via their aromatic residues.^{9,16} However, recent studies argue against the role of this motif in caveolin interactions and involvement in signalisation processes.^{8,10,12,17} Tyrosine and serine phosphorylation of cav1, as well as cysteine

palmitoylation also seem to be involved in caveolin–protein interactions. Tyrosine phosphorylation of cav1 occurs at its N-terminal domain (Tyr14) and phosphorylated cav1 serves as a scaffolding protein to recruit SH2-domain containing proteins.¹⁸ Serine phosphorylation occurs at Ser80 and is responsible for the topology change of cav1 from a membrane protein to a secreted protein.⁷ Cav1 palmitoylation has been proposed to play a role in the interaction of cav1 with lipids,¹⁶ in cholesterol binding and transport and for the assembly of signalling molecules in caveolae.^{5,6}

Caveolin-2 exists as a homo-dimer¹⁹ and acts as a scaffolding protein within caveolar membranes directly interacting with cav1 to form a stable hetero-oligomeric complex that is required for targeting to lipid rafts and for caveolae formation.²⁰ It has three isoforms: α , β and γ . The α - and β -isoforms differ in their subcellular locations and the β -isoform does not interact with cav1.²¹ Cav2 is phosphorylated at tyrosine residues 19 and 27.²² Phosphocaveolin-2 (Tyr(P)19) is localized near focal adhesions, remains associated with lipid rafts/caveolae, but no longer forms a high molecular mass hetero-oligomer with caveolin-1.^{23–25} Information about the γ -isoform of cav2 is lacking.

Caveolin 3 (also known as M-caveolin) is the muscle-specific member of caveolin family being the only member of the family that is present in striated muscle. It shares 65 % sequence identity and 85 % sequence similarity with cav1,² it also forms a large homo-oligomeric complex of 350–400 kDa molecular weight²⁶ and appears to have similar cellular functions, such as caveolae formation and cellular signalling. It is a modulator of the function of the dystrophin–glycoproteins complex with consequences for muscle diseases.²⁷

Caveolins cellular functions ascribed to caveolae are endocytosis, transcytosis and signal transduction. Caveolins also play a role in some diseases:^{26,28} cancer (they are implicated in both tumour suppression and oncogenesis), cardiovascular diseases, lung diseases and muscular dystrophy.

The structural and functional characteristics of caveolin proteins have not yet been well defined, but some useful information can be obtained through sequence based predictions using structural bioinformatics tools. In the present study, a bioinformatics investigation was performed in which, to the best of our knowledge, all the three human caveolins were considered for the first time in order to predict and compare their structural and functional properties. The sequence identity, hydrophobicity, alpha-helix and beta-turn profiles were analysed and the secondary structure elements, the disordered regions and the interacting partners were predicted. Comparing the obtained results with known structural and functional data concerning caveolin 1, the functional domains of caveolin 2 and caveolin 3 were predicted and characterized.

EXPERIMENTAL

Within this study, bioinformatics tools were employed to analyse and predict structural and physicochemical properties of the family of human caveolin proteins starting from sequence information retrieved from SwissProt data base.²⁹ Sequence analysis relating to the sequence identity of the vertebrate caveolins was performed using the BLAST tool.³⁰ The sequence similarity between the human caveolins was studied using multiple sequence analysis performed with the CLUSTALW2 on-line tool.³¹

The ProtParam tool³² was used to compute the global physicochemical properties of the caveolins and of their predicted functional regions, *i.e.*, the molecular weight, GRAVY index, net charge and aliphatic index. GRAVY is the grand average of hydropathicity and indicates the solubility of the protein. It is computed using the Kyte and Doolittle hydrophobicity scale.³³ A positive value of GRAVY reveals a hydrophobic protein and a negative value, a hydrophilic one.

The aliphatic index is a measure of the relative volume occupied by the amino acids with aliphatic side chains. This parameter is usually computed for globular proteins but it has also been used for proteins containing membrane regions.^{34,35}

The ProtScale tool³² was also used to obtain the hydrophobicity, membrane tendency and the alpha helix and beta-turn profiles for caveolins. The hydrophobicity shows the degree of hydrophobicity of each amino acid in the protein chain *versus* its position in the chain. For the hydrophobicity profiles, the Kyte and Doolittle hydrophobicity scale³² was used. Prediction of the secondary structural elements of caveolins was performed using GOR,³⁶ Jpred3,³⁷ PsiPred,³⁸ CFSSP³⁹ and ProtPred⁴⁰ on-line tools and their results were compared. Moreover, prediction of disordered regions of caveolins was obtained using ProtPred⁴⁰, RONN⁴¹ and GlobPlot⁴² tools.

Caveolins are predicted to have an unusual membrane topology. They do not cross the membrane, but they are anchored to one of its layers. A few computational tools for predicting the re-entrant loops of membrane proteins can be found: OCTOPUS⁴³ and SPOCTOPUS⁴⁴ and they were employed to predict the re-entrant loops of the caveolins.

In order to obtain the membrane tendency (MT), the transmembrane tendency indexes introduced by Zhao and London in 2006⁴⁵ were used, and for the alpha-helix and beta-turn profiles, the alpha-helix and beta-turn propensities indexes introduced by Deleage and Roux in 1987⁴⁶ were employed. The MT sequence prediction scale shows the propensity of each amino acid in the protein chain to participate in transmembrane helices⁴⁵ and the alpha-helix/beta-turn propensity is a measure of the tendency of an amino acid to adopt an alpha-helix/beta-turn structure.⁴⁶ All these profiles take into account the physicochemical background for every amino acid in the sequence, the average values of each considered property being computed using a window of 21 amino acids with the considered amino acid at its centre. The large window of 21 amino acids used for computation of profiles has been proven to be good for membrane proteins³² and the profiles that were computed in the present study with windows of 9 and 13 residues did not significantly differ from the profiles computed with a window of 21 residues (see Figs. S-1 and S-2 of the Supplementary material to this paper). In the profiles computed with a window of 21 residues, predictions for the first 10 and the last 10 amino acids of the sequence are lacking.

A short region of 8 amino acids was identified as identical in all three proteins, FEDVIAEP, and it is named as the “caveolin signature sequence” (CSS).² Therefore, for cav2 and cav3, the profiles were translated in comparison to the cav1 profile such that the caveolin



signature amino acid sequence corresponds for all caveolins: 14 positions for cav2 and 28 positions for cav3.

STRING software⁴⁷ was used to predict the interacting partners of the family of caveolin proteins. STRING is based on both experimental and predicted interaction information and it reports a network containing the highest scoring interacting partners and specifies if the interaction has been proved experimentally. As parameters setting for using STRING, a medium confidence score (*i.e.*, higher than 0.400) was chosen and the following active prediction methods: experiments, databases and text mining.

Cav1 and cav2 have two and three identified isoforms, respectively. If not specified further in the text, the cav1 and cav2 notation will refer to the alpha isoforms.

RESULTS AND DISCUSSION

No three dimensional structure of caveolin proteins have been yet determined. Homology modelling using the Geno3D tool⁴⁸ revealed that no satisfactory template could be found for their sequences (see Fig. S-3 of the Supplementary material to this paper).

The BLAST tool³⁰ revealed that caveolin amino acid sequences have been highly conserved throughout evolution, the sequence identity for vertebrates being higher than 47 %. The members of the human caveolin family share high sequence similarity: cav1 and cav2 are 58 % similar, cav1 and cav3 are 85 % similar and cav2 and cav3 are 39.0 % similar. The sequence alignment was performed using CLUSTALW2 software³¹ and the results are shown in Fig. 1, the highly conserved regions being highlighted.

Fig. 1. Sequence alignment of the alpha isoforms of cav1, cav2 and cav3; * (asterisk) indicates positions that have a single, fully conserved residue; colon indicates conservation between groups of very similar properties; period indicates conservation between groups of weakly similar properties. The single-letter amino acid code is used. The caveolin signature sequence (CSS) and the hydrophobic domain (HD) are highlighted in dark grey. The cholesterol recognition/interaction amino acid consensus (CRAC) sequences of caveolins are highlighted in light grey.



Being membrane proteins, it is expected that caveolins possess highly hydrophobic amino acid domains. The hydrophobicity profiles of the caveolin proteins obtained using the ProtScale tool³² and the Kyte & Doolittle hydrophobicity scale³³ are shown in Fig. 2, in which the hydrophobic residues have positive indexes.

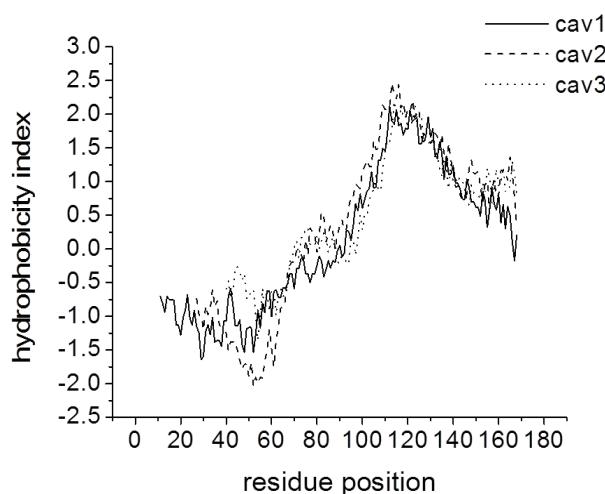


Fig. 2. Hydrophobicity profiles of cav1, cav2 and cav3 obtained using the ProtScale tool. For cav2 and cav3 proteins, the residue position was translated so as the caveolin signature sequence correspond for all caveolins.

The hydrophobicity profiles divide the amino acid sequences of caveolins in two different regions: a hydrophilic region (residues 1–93 for cav1, 1–60 for cav2 and 1–50 for cav3) and a hydrophobic region (residues 94–178 for cav1, 61–162 for cav2 and 51–151 for cav3). Figure 2 also reflects that the known hydrophobic domain of cav1 (residues 102–134) and a part of its adjacent regions are highly hydrophobic. Using this profile, it is suggested that the hydrophobic domains of cav2 and cav3 proteins are composed of the amino acid sequences 87–119 and 75–107, respectively, and this suggestion is in good agreement with the sequence alignment presented in Fig. 1.

The membrane tendency profiles were also analysed (see Fig. S-4 of the Supplementary material to this paper). The profiles are very similar for the three caveolins and, as expected, they are also similar to the hydrophobicity profiles. Moreover, both tools, OCTOPUS⁴³ and SPOCTOPUS,⁴⁴ did not predict any re-entrant loops for the caveolins.

The alpha-helix and beta-turn profiles obtained using the ProtScale tool are presented in Fig. 3.

Alpha-helix profiles show the high probability for the following amino acid sequences to form helical regions: 29–60 and 94–140 for cav1, 84–125 for cav2

and 77–141 for cav3. Except for cav1, these profiles are in reasonable agreement with the predicted membrane helices. The beta-turn profiles depict the probable regions adopting beta-structures, *i.e.*, amino acid sequences 1–38 for cav1, 20–50 for cav2 and 25–41 for cav3.

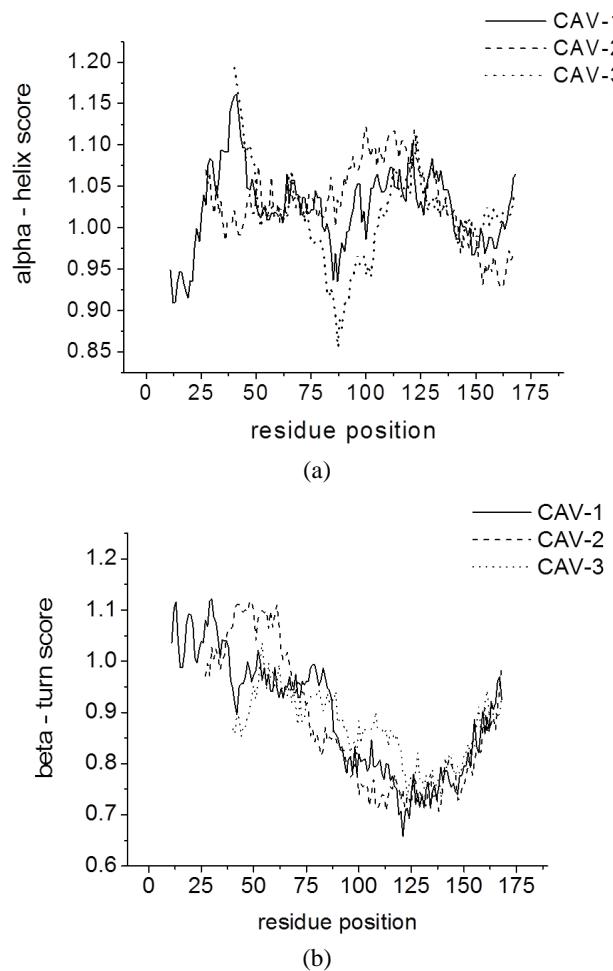


Fig. 3. Alpha-helix (a) and beta-turn (b) profile for caveolin proteins obtained using the ProtScale tool. For cav2 and cav3, the residue position was translated so that the caveolin signature sequence corresponds for all caveolins.

The alpha helix and beta-turn profiles obtained using the ProtScale tool were not fully confirmed by the prediction of the secondary structure elements using the GOR, Jpred, PsiPred, CFSSP and ProtPred tools (Fig. 4), by the predictions using RONN, GlobPlot, ProtPred of their disordered regions (Fig. 5) nor by other published data concerning their structural properties.

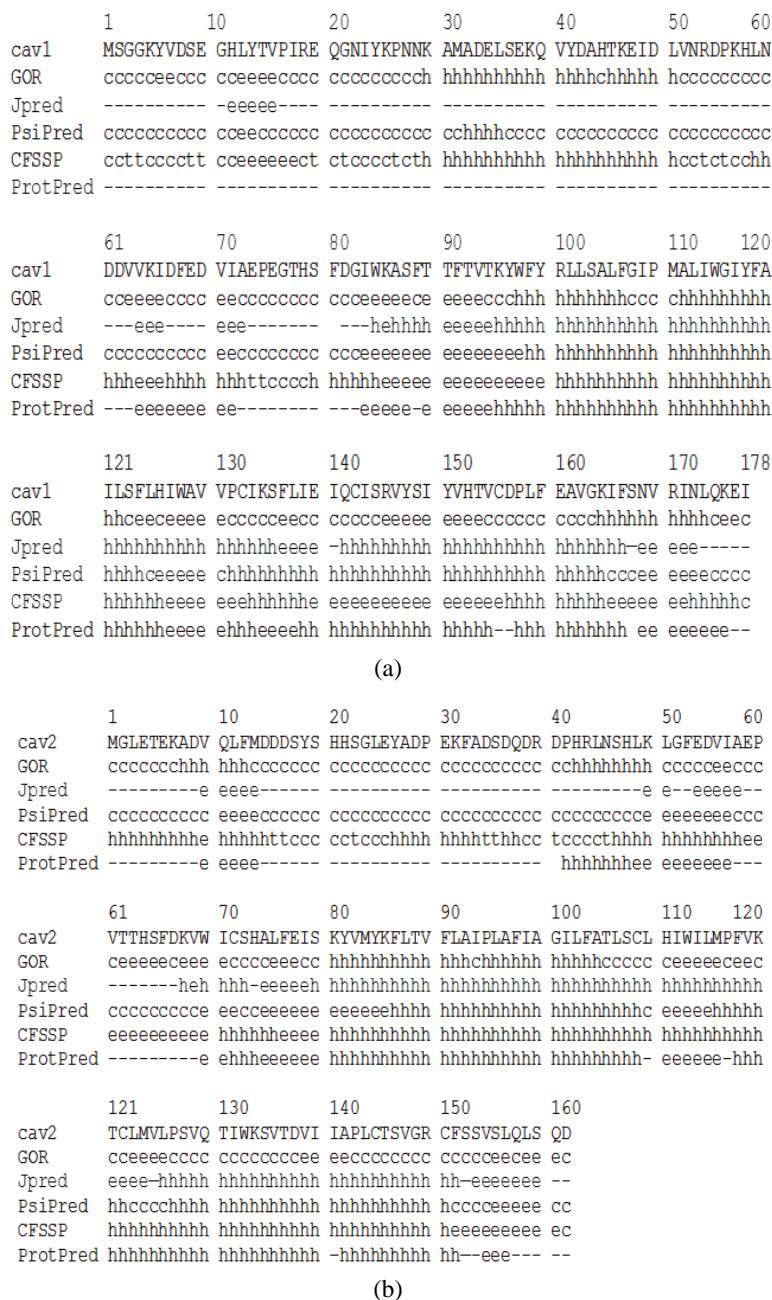


Fig. 4. Prediction of the secondary structure elements of human: a) cav1, b) cav2 and c) cav3 using the GOR, Jpred, PsiPred, CFSSP and ProtPred tools (h – helix, e – beta strand, c – coil, t – turn).

	1	10	20	30	40	50	60
cav3	MMAEEHTDLE	AQIVKDIHCK	EIDLVNRDPK	NINEDIVKVD	FEDVIAEPVG	TYSFDGWVKV	
gor	ccccccccc	hhhhhhccc	eeeeeeccc	ccccceeee	cccecccc	eeeceeeee	
jpred	-----	-----	-----	ee-	e-eee-	-----	hehhh
psipred	ccccccccc	hhhhccccc	ccccccccc	ccccceeee	cccecccc	ccccceeee	
CFSSP	hhhhhhhhh	hhhhhhhhh	hhhtctcctc	eeeeeeeeee	eeeceeeee	eehhhhhhhe	
ProtPred	-----	hhhhhh	-----	eeee	eeeeee	-----	eee
	61	70	80	90	100	110	120
cav3	SYTTFTVSKY	WCYRLLSTLL	GVPLALLWGF	LFACISFCHI	WAVVPCIKSY	LIEIQCISHI	
gor	eeeeeeccc	eeeeeeccc	cceeeeeee	eeeeccccc	eeecccc	ccccceeece	
jpred	h-eeeehhh	hhhhhhhhh	hhhhhhhhh	hhhhhhhhh	hhhhhhhee	eee-hhhhh	
psipred	eeeeeeeeee	ehhhhhhhh	hhhhhhhhh	hhhhhhhh	hhhhhhhee	eechhhhhh	hhhhhhhhh
CFSSP	eeeeeeeeee	eehhhhhhh	hhhhhhhhh	hhhhhhhhh	hhhhhhhhh	hhhhhhhhh	hhhhhhheeee
ProtPred	ee-eeeeee	hhhhhhhhh	hhhhhhhhh	hhhhhhhhh	hhhhhhhhh	-eeeeee-	eehhhhhhh
	121	130	140	150			
cav3	YSLCIRTECN	PLFAALQVCS	SIKVVLKEV				
gor	eeeccccccc	ccccccccc	eeeeeeceec				
jpred	hhhhhhhhh	hhhhhhhhh	-eeeeee-				
psipred	hhhhhhhhh	hhhhhhhhc	cceeeeeee				
CFSSP	eeeeeeeeee	hhhhhhhhh	heeeecccc				
ProtPred	hhhhhhhhh	--hhhhhhh	-eeeeeee-				

(c)

Fig. 4. (Continued) Prediction of the secondary structure elements of human: a) cav1, b) cav2 and c) cav3 using the GOR, Jpred, PsiPred, CFSSP and ProtPred tools (h – helix, e – beta strand, c – coil, t – turn).

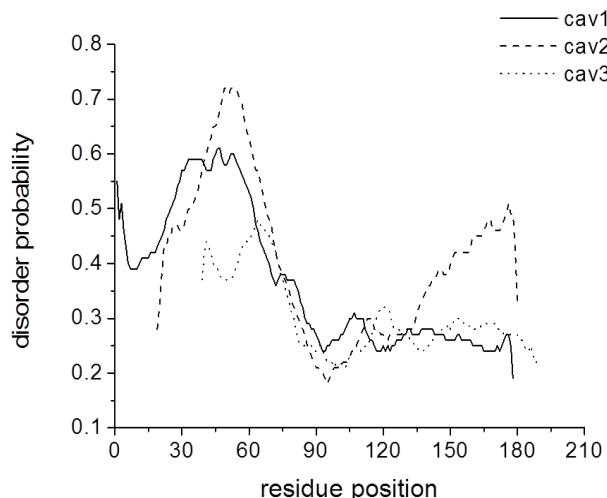


Fig. 5. Prediction of the disordered regions of caveolins using RONN software (regions with probability of disorder higher than 0.5 were predicted as disordered). For cav2 and cav3, the residue position was translated so that the caveolin signature sequence corresponds for all caveolins.

The consensus of the secondary structure predictions made using different bioinformatics tools led to the proposal that for cav1, the most probable helical regions are 98–125 (predicted by all the used tools) and 141–165 (predicted by Jpred, PsiPred and ProtPred).

Moreover, the consensus of the secondary structure predictions reflects that the 86–94 and 126–131 regions of cav1 adopt a beta-strand secondary structure. Furthermore, the prediction of the secondary structural elements of cav1 (Fig. 4) predicts that the first segment of the CSD contains a beta strand structural arrangement rich in hydrophobic amino acids and the segment 94–101 adopts a helical structure. The hydrophobic domain of cav1 is predicted to have a helical structure. The prediction of disordered regions of cav1 obtained using RONN⁴¹ reflects a quite disordered N-terminal domain with a long disordered region, 26–63 (Fig. 5).

The use of GlobPlot software⁴² for the same purpose (see Fig. S-5 of the Supplementary material to this paper) predicts the segment 24–47 as disordered and the 102–135 region as hydrophobic. ProtPred tool reflects a disordered N-terminal region (1–56) for cav1. These predictions of the secondary structure and of disordered regions for cav1 indicate that the N-terminal region of cav1 is disordered and lacks secondary structure arrangement. It was also registered that all tools used for secondary structure predictions (GOR, Jpred, PsiPred, CFSSP and ProtPred) agreed in their prediction of a helical structure for the hydrophobic domain, but concerning the prediction of disordered regions, only Jpred and ProtPred tools found the same amino acid sequences.

Comparisons of the computational and experimental data concerning the structural properties of cav1 or its fragments indicate controversial results. Based on the analysis of cav1 amino acids sequence using multiple computational algorithms, a topology model for cav1 was proposed.⁴⁹ This model proposes that the entire region of cav1 has a high probability to form helical structural elements and that the highest helical probability is found for the region 113–127. Circular dichroism (CD) and NMR spectroscopic studies of the hydrophobic domain of cav1 (96–136) revealed that this domain has a high α -helical content (57–65 %) and possess a helix-break-helix structure, the helix breaking region being 108–110.⁵⁰ Other CD studies have been performed on cav1 fragments in different environments. In particular, a CD study of cav1 fragments demonstrates that peptides 2–20 and 19–40 display 20 % content of alpha-helical structure and peptide 161–160 shows a 43 % content of alpha helix.⁵¹ Furthermore, CD analysis of cav1 (1–101) fragment in solution reflects that the amino acids 79–96 are arranged in an alpha-helix⁵². The same technique applied for the cav1 (83–102) in the presence of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) vesicles⁵³ and for cav1 (82–101) in the presence of dodecylphosphocholine (DPC) micelles⁵⁴ revealed only a partially alpha-helical structure of this isolated

domain. For a longer fragment, cav1 (82–109), NMR studies indicate alpha-helical structures for the fragments 83–88 and 93–97 of the CSD domain and a stable helical conformation only for the fragment 102–109 belonging to the first region of the hydrophobic domain.⁵⁴ CD, FTIR and NMR experimental studies of cav1 (82–109) and cav1 (83–102) in the presence of a POPC/cholesterol mixture indicated that the CSD region of cav1 contains both alpha- and beta-structures, the content of alpha-helices being higher for the longer fragment.⁵⁴ The conformational model built by Hoop *et al.*⁵⁵ proposes an anti-parallel beta-strand for the region 84–94, which was also proposed by Spisni *et al.*⁸

For cav2, there is only one helical region in alpha-helix profile, 84–125 which is in quite good agreement with predictions of the secondary structure made using Jpred, PsiPred and CFSSP (Fig. 4) and this predicted helical region also includes the predicted membrane region (87–104). The beta-turn profile indicates that the region 20–50 adopts a beta-turn structure but this was not confirmed by the Jpred algorithm. The predicted disordered regions for cav2 are the amino acid sequences 17–49 using RONN (Fig. 5) and 19–25 and 32–45 using GlobPlot. These predictions are in good agreement with the results of the GOR, Jpred and PsiPred algorithms (Fig. 4). GlobPlot software also predicts the amino acid sequence 90–119 as the membrane domain of cav2 and this prediction should be compared with the predicted helical region, amino acid sequence 84–125, with the proposed hydrophobic domain, 87–119 as well as with the predicted membrane region, 87–104. A CD experimental study of cav2 (1–73) fragment indicated a 25 % helical content of this fragment.⁵² Only the CFSSP prediction is in agreement with this experimental result.

Concerning cav3, the different algorithms used for the secondary structure predictions unanimously predict the 77–141 amino acid sequence as helical. A consensus for the membrane domain prediction was also obtained for the amino acid sequence 84–104. No disordered regions of cav3 were identified using RONN (Fig. 5) while the GlobPlot tools predict that the fragment 79–101 of cav3 is a hydrophobic domain and this prediction is in agreement with the hydrophobicity profile and with the membrane domain and secondary structure predictions. An experimental CD study revealed that fragment 1–74 for cav3 contains 25 % of α -helical secondary structure,⁵⁶ and the authors proposed that the CSD of cav3 (residues 55–72) forms an alpha-helix and that the remaining region 1–54 lacks a stable secondary structure.⁵⁷

Taking into account the sequence alignments and the analysis presented above, the following functional domains are proposed for cav2 and cav3.

For cav2: N-terminal domain – residues 1–66; CSD – residues 67–86, including the CRAC motif (residues 79–86); HD domain – residues 87–119; C-MAD domain – residues 120–135 and C-terminal domain – residues 136–162.

For cav3: N-terminal domain – residues 1–54; CSD domain – residues 55–74, including the CRAC motif (67–74); HD domain – residues 75–107; C-MAD domain – residues 108–123 and C-terminal domain – residues 124–151.

To the best of our knowledge, no published data concerning the predicted functional domains of cav2 have been published. Concerning cav3, the present predictions are in good agreement with published data of Balijepalli and Kamp, 2009.⁵⁷

The physicochemical properties of the predicted domains of cav1, cav2 and cav3 are reported in Table I.

TABLE I. Physicochemical properties of different functional domains of caveolins

Protein domain	Net charge			GRAVY			Aliphatic index		
	cav1	cav2	cav3	cav1	cav2	cav3	cav1	cav2	cav3
N-TER	-8	-11	-9	-0.858	-0.889	-0.361	74.25	60.61	95.56
CSD	+2	+1	+2	-0.265	0.160	-0.385	39.00	92.50	43.50
HD	0	0	0	2.009	2.218	2.015	171.52	171.52	165.45
C-TER	+2	+3	+3	0.212	0.774	0.796	108.46	118.89	118.21

As expected, the values presented in Table I show that the hydrophobic domains have the highest hydrophobicity reflected by GRAVY and the highest values of the aliphatic indexes. The GRAVY index of the CSD of cav2 strongly differs from the GRAVY index of the CSD of cav1 and cav3 as a result of the different amino acid contents of these domains. The amino acid sequences of cav1 and cav3 CSD differ by 6 residues while cav1 and cav2 CSD amino acid sequences differ by 12 residues. The CSD of cav2 has a higher content of hydrophobic residues: Ile71 in comparison with Lys86 in cav1 and Lys59 in cav3, Ala75 in comparison with Thr90 in cav1 and Thr63 in cav3, Leu76 in comparison with Thr91 in cav1 and Thr64 in cav3 and Val83 in comparison with Trp98 in cav1 and Trp71 in cav3 (Fig. 1). The N-terminal domain of cav2 has a more negative charge in comparison to that of cav1 and cav3. These differences could be responsible for the different folding of these proteins and/or for the different properties of the associations with themselves or with other partners.

The predicted functional partners of human cav2 and cav3 were obtained using STRING software⁴⁷ and are compared with those of caveolin-1. There are only a few common interacting partners for all the caveolins (they are presented in Table II), cav1 and cav3 having more common partners than cav1 and cav2 (Table III).

This may be due to the differences in the amino acid sequences of cav1 and cav2, but also to the lack of data concerning these interactions. Some of the predicted interacting partners of cav2 and cav3 do not contain the CBD, which is in good agreement with recent literature data.¹²



TABLE II. Proteins predicted to interact with all human caveolins: cav1, cav2 and cav3 (the interactions that are experimentally proven are marked with an asterisk)

Protein	Interaction with			CBM and/or CBM-like motif containing molecule ¹⁷
	cav1	cav2	cav3	
v-src Sarcoma viral oncogene homolog	X*	X*	X	—
FYN oncogene related to SRC	X*	X	X	—
Nitric oxide synthase 3 (NOS3)	X*	X	X	X
Gap junction protein, alpha 1 (GJA1)	X*	X	X	X
Flotillin 2 (FLOT2)	X*	X*	X*	X
Integrin, beta 4 (ITGB4)	X*	X	X	—
Integrin, beta 5 (ITGB5)	X*	X	X	—
Integrin, beta 6 (ITGB6)	X*	X	X	—
Integrin, beta 7 (ITGB7)	X*	X	X	—
Integrin, beta 3 (ITGB3)	X*	X	X	—
Integrin, beta 8 (ITGB8)	X*	X	X	—
Integrin, beta 1 (ITGB1)	X*	X	X	—

TABLE III. Proteins predicted to interact with cav1 and cav2, respectively with cav1 and cav3 (the experimentally proven interactions are marked with an asterisk)

Protein	Interaction with			CBM and/or CBM-like motif containing molecule ¹⁷
	cav1	cav2	cav3	
Caveolin 2 (CAV2)	X*	—	—	—
Insulin receptor (INSR)	X*	—	X	X
Insulin receptor substrate 1 (ISR1)	X*	—	X	X
Transient receptor potential cation channel, subfamily C	X*	—	X	X
Nitric oxide synthase 3 (NOS3)	X*	X	X	X
Ras-related C3 botulinum toxin substrate 1 (RAC1)	X*	—	X	X
Dystroglycan 1(DAG1)	X*	—	X*	X
Polymerase I and transcript release factor (PTRF)	X*	—	X	—
Phospholipase D1, phosphatidylcholine-specific	X*	—	X	X
Mal (MALL)	X*	X*		X
Platelet-derived growth factor receptor	X*	—	X*	X
Phospholipase D2 (PLD2)	X*	X*		X
Ras homolog gene family, member A (RHOA)	X*	—	X	—
Nitric oxide synthase 2, inducible (NOS2)	X*	—	X*	X
V-Akt murine thymoma viral oncogene homolog 1	X	—	X	—
Calcium channel, voltage-dependent, L type	X	—	X	—
Nitric oxide synthase 1 (NOS1)	X*	—	X*	X
Calcium channel, voltage-dependent, N type	X	X		X
Solute carrier family 2 (SLC2A4)	X	—	X	—
RAS p21 protein activator (RASA1)	X*	X	—	—



TABLE III. Continued

Protein	Interaction with			CBM and/or CBM-like motif containing molecule ¹⁷
	cav1	cav2	cav3	
Syndecan 2 (SDC2)	—	X	—	—
V-Ha-ras Harvey rat sarcoma viral oncogene homolog	X*	X	—	—
Solute carrier family 8, member 1 (SLC8A1)	X*	—	X*	X
Synaptosomal-associated protein (SNAP23)	X*	—	X*	X
SMAD family member 3 (SMAD3)	X*	—	X*	X
Stromal interaction molecule 1 (STIM1)	X*	—	X*	—
Dopamine receptor D1(DRD1)	X*	X*	—	X

CONCLUSIONS

For the first time, a bioinformatics study comparing all the three amino acid sequences of human caveolins is presented. Taking into account the known data about cav1 (the most studied from this family), the functional domains of cav2 and cav3 were predicted and characterised. The predicted functional domains of cav2 are: N-terminal domain – residues 1–66; CSD – residues 67–86, including the CRAC motif (residues 79–86); HD domain – residues 87–119; C-MAD domain – residues 120–135 and C-terminal domain – residues 136–162. Similarly, the predicted functional domains for cav3 are: N-terminal domain – residues 1–54; CDS domain – residues 55–74, including the CRAC motif (67–74); HD domain – residues 75–107; C-MAD domain – residues 108–123 and C-terminal domain residues 124–151.

The amino acid sequence of cav2 differs slightly from that of cav1 and cav3 and its N-terminal is more hydrophilic and has a higher overall negative charge than the N-terminal domains of cav1 and cav3. All the caveolins possess three cysteine residues in the C-terminal domains, their positions being conserved for cav1 and cav3 but not for cav2. Therefore, for cav3, the three cysteine residues are most probably palmitoylated. The oligomerization domain of cav1 contains 41 residues (region 61–101). Sequence alignment revealed that in the homology region of cav3 (residues 34–74), there are only 7 distinct residues (17 %) and in the homology region of cav2 (residues 46–86), there are 17 distinct residues (41 %). This sequence dissimilarity may be responsible for the difference in oligomerization of the three caveolins: cav1 and cav3 being able to form high weight oligomers while cav2 forms only homo-dimers.¹⁹

The CSD of the three caveolins also differ in sequence: there are 6 distinct residues between the amino acids sequences of cav1 (82–101) and cav3 (55–74), and 12 distinct residues between the amino acids sequences of cav1 (82–101) and cav2 (67–86). These differences may be responsible for the specificity of the interactions of the three caveolins, as the CSD domain has been proposed to be involved in interactions with functional partners.



The amino acid sequences of the hydrophobic domains of the three caveolins are the most similar in sequence, they are highly hydrophobic and are predicted to comprise helical structures. This expected result is in good agreement with experimental data concerning the cav1 (96–136) fragment in the presence of lyso-myristoylphosphatidylglycerol.⁵⁰

The amino acid sequences of the C-terminal domains of the caveolins are the most divergent: there are 18 distinct residues between the amino acid sequences of cav1 (135–178) and cav3 (108–151) and 32 distinct residues between the amino acid sequences of cav1 (135–178) and of cav2 (120–162). The amino acid sequence of the C-terminal segment of cav1 is mostly hydrophilic, and has two proposed functions: membrane anchoring for its first part (C-MAD, residues 135–150) and protein–protein interactions for its last part (residues 168–178), which is also involved in homotypic interactions⁵⁸ and contributes to the side-by-side clustering of homo-oligomers of caveolin-1.⁷ There are few experimental results concerning the C-terminal region of cav2, but it is known that the Ser154–Val155–Ser156 motif of the C-terminal of cav2 is essential for insulin-induced phosphorylation and nuclear targeting of extracellular signal-regulated kinase;^{59,60} this motif being absent in the sequences of cav1 and cav3. In addition, experimental data suggest that the C-terminal domain of cav2 is necessary for targeting cav2 to caveolae.²¹ Therefore, the divergence of the amino acid sequences of the C-terminal regions of the three caveolins may also contribute to the different location and distinct interacting partners of these proteins.

The performed bioinformatics analysis of the amino acid sequences of human caveolins highlighted both similar and distinct structural and functional properties of caveolins in correlation with their amino acid sequences. It also highlighted the limitations of the currently existing tools concerning the predictions of structural features of membrane proteins. An agreement between the different experimental structural data, especially obtained for cav1, is lacking. This reflects that the structure adopted by the different regions of caveolin may be distinct as a result of the different lengths of the considered fragments and of the interactions with their environments. It must also be taken into account that the computational tools for the prediction of the structural features of membrane proteins have limitations.

Further research must be realised to elucidate the interaction network of caveolins with physiological implications. From this point of view, knowledge concerning their full length spatial structures would substantially contribute to an understanding of their functions.

SUPPLEMENTARY MATERIAL

Figures S-1–S-5 are available electronically from <http://www.shd.org.rs/JSCS/>, or from the corresponding author on request.

Acknowledgement. This work is a result of a bilateral collaboration Romania – France project, Programme Brancusi, 492/2011.

ИЗВОД

БИОИНФОРМАТИЧКО ИСПИТИВАЊЕ СТРУКТУРЕ И ФУНКЦИОНАЛНИХ СВОЈСТАВА ХУМАНИХ ПРОТЕИНА КАВЕОЛИНА

ADRIANA ISVORAN^{1,2}, DANA CRACIUN³, ALECU CIORSAC⁴, NAHUEL PERROT⁵, VERONICA BESWICK^{5,6}, PIERRE NEDELLEC⁵, ALAIN SANSON⁵ и NADEGE JAMIN⁵

¹Department of Biology–Chemistry, West University of Timisoara, 16 Pestalozzi, 300316 Timisoara, Romania, ²Laboratory of Advanced Researches in Environmental Protection, West University of Timisoara, 4 Oituz, 300086 Timisoara, Romania, ³Teacher training Department, West University of Timisoara, 4 V.Pirvan, 300223 Timisoara, Romania, ⁴Department of Physical Education and Sport, Politehnica University of Timisoara, 2 P-ta Victoriei, 300306 Timisoara, Romania, ⁵Commissariat à l'Energie Atomique (CEA), Institute of Biology and Technologies of Saclay (iBiTecS) 91191 Gif-sur-Yvette Cedex, France и ⁶Department of Physics, Université d'Evry-val-d'Essonne, 91025 Evry Cedex, France

Урађена је биоинформатичка студија у циљу предвиђања и упоређивања структуре и функционалних својстава хуманог кавеолина-1, -2 и -3. Компјутерски процењена физичко–хемијска својства, секундарна структура и интерагујући партнери кавеолина-2 и -3 су упоређивани са експериментално доказаном структуром и својствима кавеолина-1. Ови подаци, заједно са примарном секвенцијом три кавеолина, омогућили су предвиђање и карактеризацију функционалних домена кавеолина-2 и -3. Хидрофобни региони ова три протеина имају веома сличну секвенцију и физичко–хемијска својства, што је у складу са њиховом мембрanskом локализацијом и функцијама. Највећа разлика у секвенцији и својствима нађена је у С-терминалном региону кавеолина, те се може претпоставити да је он одговоран за различите интеракције које би имале директне последице на пренос сигнала.

(Примљено 16 јула, ревидирано 21. септембра 2013)

REFERENCES

1. A. Schlegel, D. Volonte, J. A. Engelmann, F. Galbiati, P. Mehta, X. L. Zhang, P. E. Scherer, M. P. Lisanti, *Cell. Signal.* **10** (1998) 457
2. A. W. Cohen, R. Hnasko, W. Schubert M. P. Lisanti, *Physiol. Rev.* **84** (2004) 134
3. P. E. Scherer, Z. Tang , M. Chun, M. Sargiacomo, H. F. Lodish, M. P. Lisanti, *J. Biol. Chem.* **270** (1995) 16395
4. H. Kogo, T. Fujimoto, *FEBS Lett.* **465** (2000) 119
5. A. Uittenbogaard, E. J. Smart, *J. Biol. Chem.* **275** (2000) 25595
6. M. O. Parat, P. L. Fox, *J. Biol. Chem.* **276** (2001) 15776
7. A. Schlegel, M. P. Lisanti, *J. Biol. Chem.* **275** (2000) 21605
8. E. Spisni, V. Tomasi, A. Cestaro, S. C. E. Tosatto, *Biochem. Biophys. Res. Comm.* **338** (2005) 1383
9. A. Arbuzova, J. Wang, L. Wang, G. Hangyas-Mihalyne, D. Murray, B. Honig, S. McLaughlin, *Biochemistry* **39** (2000) 10330
10. J. Couet, S. Li, T. Okamoto, T. Ikezu, M. P. Lisanti, *J. Biol. Chem.* **272** (1997) 6525
11. J. Couet, M. Sargiacomo, M. P. Lisanti, *J. Biol. Chem.* **272** (1997) 30429
12. B. M. Collins, M. J. Davis, J. F. Hancock, R. G. Parton, *Dev. Cell.* **23** (2012) 11
13. H. Li, V. Papadopoulos, *Endocrinology* **139** (1998) 4991



14. C. Le Lan, J. Gallay, M. Vincent, J. M. Neumann, B. de Foresta, N. Jamin, *Eur. Biophys. J.* **39** (2010) 307
15. M. Sáinz-Jaspeado, J. Martin-Liberal, L. Lagares-Tena, S. Mateo-Lozano, X. G. del Muro, O. M. Tirado, *Oncotarget* **2** (2011) 305
16. R. G. W. Anderson, *Annu. Rev. Biochem.* **67** (1998) 199
17. D. P. Byrne, C. Dart, D. J. Rigden, *PLOS ONE* **7** (2012) e44879
18. H. Lee, D. Volonte, F. Galbiati, P. Iyengar, D. M. Lublin, D. B. Bregman, M. T. Wilson, R. Campos-Gonzales, B. Bouzahzah, R. G. Pestell, P. E. Scherer, M. P. Lisanti, *Mol. Endocrinol.* **14** (2000) 1750
19. P. E. Scherer, T. Okamoto, M. Chun, I. Nishimoto, H. F. Lodish, M. P. Lisanti, *Proc. Natl. Acad. Sci. U.S.A.* **93** (1996) 131
20. P. E. Scherer, R. Y. Lewis, D. Volonte, J. A. Engelmann, F. Galbiati, J. Couet, D. S. Kohtz, E. van Donselaar, P. Peters, M. P. Lisanti, *J. Biol. Chem.* **272** (1997) 29337
21. H. Kogo, K. Ishiguro, S. Kuwaki, T. Fujimoto, *Arch. Biochem. Biophys.* **401** (2002) 108
22. H. Lee, D. S. Park, X. B. Wang, P. E. Scherer, P. E. Schwartz, M. P. Lisanti, *J. Biol. Chem.* **277** (2002) 34556
23. X. B. Wang, H. Lee, F. Capozza, S. Marmon, F. Sotgia, J. W. Brooks, R. Campos-Gonzalez, M. P. Lisanti, *Biochemistry* **43** (2004) 13694
24. G. Sowa, M. Pypaert, D. Fulton, W. C. Sessa, *Proc. Natl. Acad. Sci. U.S.A.* **100** (2003) 6511
25. G. Sowa, L. Xie, L. Xu, W. C. Sessa, *Biochemistry* **47** (2008) 101
26. T. M. Williams, M. P. Lisanti, *Ann. Med.* **36** (2004) 584
27. K. S. Song, P. E. Scherer, Z. Tang, T. Okamoto, S. Li, M. Chafel, C. Chu, D. S. Kohtz, M. P. Lisanti, *J. Biol. Chem.* **271** (1996) 15160
28. M. Shatz, M. Liscovitch, *Int. J. Radiat. Biol.* **84** (2008) 177
29. R. Apweiler, A. Bairoch, C. H. Wu, *Curr. Opin. Chem. Biol.* **8** (2004) 76
30. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **215** (1990) 403
31. M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGgettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilim, R. Lopez, J. D. Thompson, T. J. Gibson, D. G. Higgins, *Bioinformatics* **23** (2007) 2947
32. E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, A. Bairoch, in *The Proteomics Protocols Handbook*, J. M. Walker, Ed., Humana Press, Totova, NJ, 2005, p. 571
33. J. Kyte, R. F. Doolittle, *J. Mol. Biol.* **157** (1982) 105
34. B. Vaseeharan, S. Valli, *J. Proteomics Bioinform.* **4** (2011) 1
35. R. Mohan, S. Venugopal, *Bioinformation* **8** (2012) 722
36. J. Garnier, J. F Gibrat, B. Robson, in *Methods in Enzymology*, Vol. 266, R. F. Doolittle Ed., Academic Press, Lemoyne, PN, 1996, p. 540
37. C. Cole, J. D. Barber, G. J. Barton, *Nucl. Acids Res.* **36** (2008) W197
38. D. W. Buchan, S. M. Ward, A. E. Lobley, T. C. Nugent, K. Bryson, D. T. Jones, *Nucl. Acids Res.* **38** Suppl. (2010) W563
39. P. Y. Chou, G. D. Fasman, *Biochemistry* **13** (1974) 211
40. B. Rost, G. Yachdav, J. Liu, *Nucl. Acids Res.* **32** (2004) W321
41. Z. R. Yang, R. Thomson, P. McNeil, R. M. Esnouf, *Bioinformatics* **21** (2005) 3369
42. R. Linding, R. B. Russell, V. Neduvia, T. J. Gibson, *Nucl. Acids Res.* **31** (2003) 3701
43. H. Viklund, A. Elofsson, *Bioinformatics* **24** (2008) 1662
44. H. Viklund, A. Bernsel, M. Skwark, A. Elofsson, *Bioinformatics* **24** (2008) 2928



45. G. Zhao, E. London, *Protein Sci.* **15** (2006) 1987
46. G. Deleage, B. Roux, *Protein Eng.* **1** (1987) 2890
47. D. Szklarczyk, A. Franceschini, M. Kuhn, M. Simonovic, A. Roth, P. Minguez, T. Doerks, M. Stark, J. Muller, P. Bork, L. J. Jensen, C. von Mering, *Nucl. Acids Res.* **39** (2011) D561
48. C. Combet, M. Jambon, G. Deleage, C. Geourjon, *Bioinformatics* **18** (2002) 213
49. R. G. Parton, M. Hanzal-Bayer, J. F. Hancock, *J. Cell. Sci.* **119** (2006) 787
50. J. Lee, K. J. Glover, *Biochim. Biophys. Acta* **1818** (2012) 1158
51. M. E. Schroeder, H. A. Hostetler, F. Schroeder, J. M. Ball, *J. Amino Acids* (2012) ID 575180
52. I. Fernandez, Y. Ying, J. Albanesi, R. G. Anderson, *Proc. Natl. Acad. Sci. U.S.A.* **99** (2002) 11193
53. R. M. Epand, B. G. Sayer, R. F. Epand, *J. Mol. Biol.* **345** (2005) 339
54. C. Le Lan, J. M. Neumann, N. Jamin, *FEBS Lett.* **580** (2006) 5301
55. C. D. Hoop, V. N. Sivanandam, R. Kodaly, M. N. Srnec, P. C. A. van der Wel, *Biochemistry* **51** (2012) 90
56. S. R. Fuhs, P. A. Insel, *J. Biol. Chem.* **286** (2011) 14830
57. R. C. Balijepalli, T. J. Kamp, *Prog. Biophys. Mol. Biol.* **98** (2009) 149
58. K. S. Song, Z. Tang, S. Li, M. P. Lisanti, *J. Biol. Chem.* **272** (1997) 4398
59. H. Kwon, K. Jeong, E. M. Hwang, J. Y. Park, Y. Pak, *J. Cell. Mol. Med.* **15** (2011) 888
60. K. Jeong, H. Kwon, J. Lee, D. Jang, E. M. Hwang, J. Y. Park, Y. Pak, *Traffic* **13** (2012) 1218.