ACADEMIC PRESS

Biochemical and Biophysical Research Communications 296 (2002) 156–160

BBRC

www.academicpress.com

Structural characterization of the second TSP1-module of human thrombospondin☆

Emőke Roszmusz,[a] András Patthy,[b] Mária Trexler,[a] and László Patthy[a,*]

[a] *Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, H-1113, Karolina ut 29, Budapest, P.O. Box 7, H-1518, Hungary*

[b] *Agricultural Biotechnology Center, Gödöllő, P.O. Box 170, H-2100, Hungary*

Received 28 June 2002

Abstract

The TSP1-module has been first identified as the type 1 repeat of thrombospondin-1. Members of this extracellular module-family have since been shown to be present in several hundred metazoan proteins as well as in proteins of some protists. Despite the widespread occurrence and biological importance of this module-type, relatively little is known about their three-dimensional structure. To define the structural features of this important module-family, we have expressed the second TSP1-domain of human thrombospondin 1 in *Escherichia coli.* Amino acid sequencing of proteolytic fragments of the recombinant protein have shown that its disulfide bonds connect the six conserved cysteines in a 1–5, 2–6, 3–4 pattern. Circular dichroism studies on the recombinant protein indicate that the disulfide-bonded TSP1-module consists primarily of distorted β-strands.

Keywords: Circular dichroism spectroscopy; Disulfide connectivity; Midkine; Pleiotrophin; Properdin; Thrombospondin

Thrombospondin-1 is a large extracellular multidomain protein that mediates cell-to-cell and cell-to-matrix interactions. The protein was shown to contain three types of repeats: the first—the TSP1 module—was found to be homologous with a circumsporozoite protein from *Plasmodium falciparum*, the second type is homologous with epidermal growth factor precursor, and the third type showed homology with parvalbumin and calmodulin [1].

The three TSP1 repeats of thrombospondin-1 were later shown to display significant homology with domains of several complement proteins [2] and related domains were subsequently identified in a large variety of metazoan proteins. In the latest release of the SMART database ([3]; http://smart.embl-heidelberg.de/) in a nonredundant sequence database there were 1175 TSP1 domains in 509 metazoan or protist proteins.

Through its involvement in cell–matrix interactions thrombospondin-1 influences cellular phenotype and the structure of the extracellular matrix, and thus plays a crucial role in tissue remodeling processes such as angiogenesis and neoplasia. Recent studies have shown that thrombospondin-1 is an endogenous inhibitor of angiogenesis and tumor growth (for a review see [4]). Structure–function studies on thrombospondin-1 have revealed that the angiostatic action of this protein is mediated by its type 1 repeats [5]. Bein and Simons [6] have recently provided evidence that the first and second TSP1 domains of thrombospondin type 1 interact with matrix metalloproteinase 2 and may regulate metalloproteinase activity, raising the possibility that this interaction may be relevant to the angiostatic action of TSP1 domains. It is interesting in this respect that the brain-specific p53-target gene, BAI1, containing five thrombospondin type 1 repeats was also found to inhibit experimental angiogenesis. A recombinant protein corresponding to the TSP-type 1 repeats of this gene product inhibited in vivo neovascularization [7].

☆ *Abbreviations:* HPLC, high pressure liquid chromatograpy; PCR, polymerase chain reaction; PTH, phenylthiohydantoin; SDS–PAGE, SDS–polyacrylamide gel electrophoresis.

* Corresponding author. Fax: +361-4665-465.
E-mail address: patthy@enzim.hu (L. Patthy).

0006-291X/02/$ - see front matter
PII: S0006-291X(02)00826-4

Despite the obvious biological importance of TSP1-modules, little is known about their three-dimensional structure. Neutron and X-ray scattering studies on the human complement protein properdin provided the most relevant information on the three-dimensional structure of the thrombospondin type 1 repeats [8]. The mean dimensions of the TSP1 domains of properdin were determined to be approximately 4 nm × 1.7 nm × 1.7 nm, indicating that they have a rather elongated structure.

Recently, it has been suggested that heparin-binding growth-associated molecule HB-GAM and other members of the pleiotrophin/midkine family consists of two tandem TSP1-related domains [9]. NMR spectroscopy of midkine and HB-GAM have shown that each of their domains consist of three antiparallel β-strands [9,10], suggesting a related structure for TSP1 domains in general. However, the domains of midkines/pleiotrophins show only very low sequence similarity with typical TSP1 domains, therefore the reliability of this prediction is unclear. To determine the key structural features of typical TSP1 modules we have expressed the second TSP1-domain of human thrombospondin-1 in *Escherichia coli*, determined its disulfide-bond connectivity, and with the aid of circular dichroism studies have predicted its secondary structural elements.

Experimental procedures

Materials and methods

Restriction enzymes, PCR primers, vectors, and bacterial strains. Restriction enzymes were purchased from Promega (Madison, WI, USA) and New England Biolabs (Beverly, MA, USA). The M13 sequencing reagents used for dideoxy sequencing of cloned DNA fragments were from Amersham Pharmacia Biotech (Uppsala, Sweden). PCR primers were obtained from Integrated DNA Technologies (Coralville, IA, USA) and from Pharmacia Biotech (Vienna, Austria).

Plasmid pMed23 [11] was from Dr. P. Venetianer (Biological Research Center, Szeged, Hungary). *E. coli* strain JM-109 was used to propagate and amplify expression plasmids. The pMed23 expression plasmid contains an ampicillin resistance gene for the selection of the positive clones.

Cloning and Expression of the second TSP1 module of human thrombospondin-1. The DNA segment coding for the second TSP1-module of human thrombospondin-1 protein (residues Gln^{434}– Pro^{490}) was amplified with the 5′ GAC GAA TTC TCG AGC AGG ATG GTG GCT GG 3′ sense, and 5′ GCG GTC GAC TCA AGA TCT GGG GCA GGC GTC TTT C 3′ antisense primers from a human genomic DNA. The amplified DNA was digested with *Eco*RI and *Sal*I restriction endonucleases and ligated into M13mp19 Rf digested with the same enzymes. The sequence of the cloned DNA was determined by dideoxy sequencing on both strands.

The DNA fragment encoding the second TSP1-module of thrombospondin-1 was excised from M13mp19 with *Eco*RI–*Hin*dIII digestion and ligated into pMed23 expression vector cut with *Pvu*II–*Hin*dIII. The *Eco*RI end of the digested DNA fragment and the *Pvu*II end of the digested vector were filled with Klenow polymerase. *E. coli* JM-109 cells were transformed with the ligation mixture and plated on LB medium (1% tryptone, 0.5% yeast extract, and 1% NaCl) containing 100 μg/ml ampicillin.

Escherichia coli JM-109 cells carrying the expression vector were grown, and expression of β-galactosidase fusion proteins was induced with 100 μM isopropyl β-D-thiogalactopiranoside. The fusion products were isolated from inclusion bodies and the isolated recombinant proteins were refolded by dialysis against 100 mM Tris and 10 mM EDTA, pH 8.0 buffer, for 24 h, and later against 0.1 M ammonium bicarbonate pH 8.0 buffer.

The β-galactosidase moiety was removed by limited trypsin digestion. The recombinant protein (1 mg/ml) was dissolved in 0.1 M ammonium bicarbonate buffer and incubated with 1 μg/ml trypsin (Sigma, St. Louis, MO, USA) at 25 °C for 15 min. Reaction was arrested with 2 mM phenylmethanesulfonyl fluoride (Serva, Heidelberg) and lyophilized. The digested recombinant protein was separated from the β-galactosidase fragment on Sephadex G-50 column, equilibrated with 0.1 M ammonium bicarbonate, pH 8.0. Fractions containing the TSP1-module were pooled, and lyophilized. SDS–PAGE analysis of the trypsin-digested protein indicated a 4-kDa reduction of the molecular mass compared with that of the β-galactosidase-fusion protein. Sequence analysis with a PE-Applied Biosystems Procise protein sequencing system showed that the trypsin cleavage occurred at the boundary of the β-galactosidase region of the β-gal fusion protein. The amino acid sequence of the resulting protein is

ILE**QDGGWSHWSPWSSCSVTCGDGVITRIRLCNSPSPQMN GKPCEGEARETKACKKDACP**RS

where the residues corresponding to the second TSP1 domain of human thrombospondin are in bold. Residues ILE. . .RS come from the vector construct.

Protein analyses. The composition of protein samples was analyzed by SDS–PAGE using 11–22% linear polyacrylamide gradient slab gels under both reducing and nonreducing conditions [12]. The gels were stained with Coomassie brilliant blue G-250. The concentration of the recombinant TSP1-module was determined using the extinction coefficient 17,250 $M^{-1}\,cm^{-1}$, calculated according to a described procedure [13].

Circular dichroism spectroscopy. CD spectra were measured over the range of 190–250 nm by using a JASCO J-720 spectropolarimeter thermostatted with a Neslab RT-111 water bath. The measurements were carried out in 1 mm pathlength cells and protein solutions of approximately 0.1–0.3 mg/ml in 10 mM Tris–HCl, pH 8.0 buffer. All spectra were measured at 25 °C with a 16 s time constant and a scan rate of 10 nm/min. The spectral slit width was 1.0 nm. All measurements represent the computer average of three scans. Secondary structure of the recombinant protein was estimated from the CD spectra with the CDPro software ([14–16] http://lamar.ColoState.EDU/∼sreeram/CDPro/index.html).

Secondary structure prediction. Secondary structure prediction of the second TSP1-module of human thrombospondin 1, based on multiple alignments of TSP1 modules, was carried out with described procedures [17–19] using the PHD program of the Predict Protein server (http://www.embl-heidelberg.de/predictprotein/submit_def.html). This server predicts secondary structural elements by evaluating the relative probabilities that a given segment can be assigned to helix, strand, or loop. The estimated accuracy of this multiple alignment-based method for the correct prediction of secondary elements is about 72%.

Localization of disulfide bonds. One hundred and ten microgram samples of HPLC purified recombinant TSP1 (dissolved in 110 μl 0.1 M ammonium acetate, pH 5.2) were digested with 5.5 μg chymotrypsin (Sigma, St. Louis, MO, USA) at 37 °C for 18 h. Samples were analyzed by reverse-phase HPLC on an Aquapore OD300 (220× 2.1 mm) column (PE-Applied Biosystems) in 0.1% (by volume) trifluoroacetic acid with a linear gradient of acetonitrile. N-terminal amino

acid sequencing of the resulting peptides was performed on a PE-Applied Biosystems. Procise protein sequencing system.

Results and discussion

Circular dicroism spectra of the TSP1-module show two distinct characteristic maxima at 212 and 229 nm and a trough at 220 nm (Fig. 1). Analysis of the spectra with the CDPro software predicted 0.923 distorted β-strand, 0.043 regular β-strand, 0.034 turn but no regular or distorted α-helix. This conclusion is in harmony with the results of secondary structure predictions based on the multiple alignment of TSP1-modules, inasmuch as it also indicated the absence of alpha helices and very little regular β-strand (Fig. 2). Thermal unfolding of the TSP1 module, monitored by changes in its CD spectra at 229 nm, has indicated that the domain collapses with a single thermal transition (T_m value of 52 °C) suggesting that the sample is structurally homogenous and corresponds to a single protein-fold (Fig. 3).

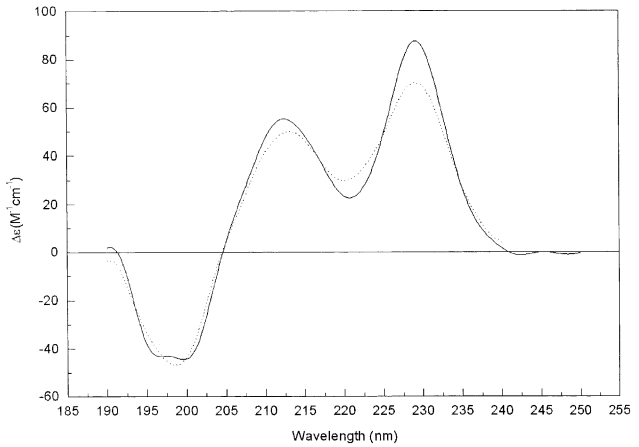

Fig. 1. Far UV circular dichroism spectra of recombinant TSP1-module of human thrombospondin 1. The solid line indicates the spectrum of the recombinant TSP1-module, the dotted line indicates the CDPro-predicted spectrum of a protein consisting of 0.923 distorted β-strand, 0.043 regular β-strand, 0.034 turn, 0.000 regular helix, 0.000 distorted helix, and 0.000 unordered structure. Spectra were obtained in 10 mM Tris–HCl, pH 8.0, at 25 °C using 0.1 mg/ml of protein.

The CD spectrum of the second TSP1-module of thrombospondin is quite similar to that of properdin (a protein consisting mainly of six tandem TSP1 modules) which also shows maxima of positive ellipticity at 213 and 231 nm and a trough at 220 nm [20]. It seems thus likely that such a CD spectrum may be characteristic of TSP1-domains.

It must be pointed out, however, that this conclusion conflicts with the results obtained on individual TSP1 repeats of properdin expressed in *E. coli* and refolded through a denaturation–renaturation cycle [21]. The CD spectra of these recombinant TSP1 domains lack the positive ellipticity maxima observed on native properdin. One possible explanation for this difference is that the recombinant TSP1 domains did not fully adopt the native fold during the denaturation–renaturation cycle. The fact that recombinant TSP1 domains of properdin failed to bind to either C3b or sulfatide, whereas polyclonal antibodies raised against the fifth TSP1 module inhibited binding of native human properdin to solid-phase C3b or sulfatides may also be due to improper refolding of recombinant proteins.

To further clarify structural characteristics of TSP1-modules, we have determined the disulfide bond pattern of the second TSP1-domain of human thrombospondin-1. Digestion of the recombinant protein with chymotrypsin yielded two fragments (Fragments A and B in Table 1) that contained all six cysteine residues. Sequence analysis of the chymotryptic peptides constituting these fragments allowed unambiguous assignment of all disulfide linkages. Fragment A contained a chymotryptic peptide corresponding to residues Cys32–Thr51 of the recombinant protein, establishing that one of the disulfide-bonds connects residues Cys32 and Cys44 (Table 1).

```
cys          ............C...C..........C..........C.........C.....C.
phd          .....EE..............EEEEEE.....................................
tshup1_1     ADDGWSPWSEWTSCSTSCGNGIQQRGRSCDSL.....NNRCEGSSVQTRTCHIQE.CD
tshup1_2     QDGGWSHWSPWSSCSVTCGDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDA.CP
tshup1_3     INGGWGPWSPWDICSVTCGGGVQKRSRLCNNPTPQFGGKDCVGDVTENQICNKQD.CP
prop_human_4 VAGGWGPWGPVSPCPVTCGLGQTMEQRTCNHPVPQHGGPFCAGDATRTHICNTAVPCP
csp_plaba    KQIRDSITEEWSQCNVTCGSGIRVRKRKGSNKKAE....DLTLEDIDTEICKMD.KCS
bai11        ATGGWKLWSLWGECTRDCGGGLQTRTRTC.LPAPGVEGGGCEGVLEEGRQCNREA.CG
bai12        GDPAAEEWSPWSVCSSTCGEGWQTRTRFCVSSS...YSTQCSGPLREQRLCNNSAVCP
bai13        VHGAWDEWSPWSLCSSTCGRGFRDRTRTCR..PPQFGGNPCEGPEKQTKFCNIAL.CP
bai14        VDGNWNEWSSWSACSASCSQGRQQRTRECNG..PSYGGAECQGHWVETRDCFLQQ.CP
bai15        VDGKWQAWASWGSCSVTCGAGSQRRERVCSG..PFFGGAACQGPQDEYRQCGTQR.CP
co6_human_c  VDGQWGCWSSWSTCD...ATYKRSRTRECNNPAPQRGGKRCEGEKRQEEDCTFSIME.
co7_human_c  VDGGWSCWSSWSPCV...QG.KKTRSRECNNPPPSGGGRSCVGETTESTQCEDEELE.
co8a_human_c ADGSWSCWSSWSVCR...AG.IQERRRECDNPAPQNGGASCPGRKVQTQAC.......
co8b_human_c IDGKWNCWSNWSSCS...GR.RKTRQRQCNNPPPQNGGSPCSGPASETLDCS......
```

Fig. 2. Multiple alignment of representative members of the TSP1 module family. The top line highlights the six conserved cysteines of TSP1-modules, PHD indicates the secondary structural elements (E for β-strand) predicted for the second TSP1 module of human thrompospondin 1 with the PHD program. *Abbreviations*: tshup1_1, tshup1_2, tshup1_3—the first, second, and third TSP1 modules of human thrombospondin 1; prop_human_4—the fourth TSP1 module of human properdin; csp_plaba—the TSP1 module of the circumsporozoite protein of *Plasmodium berghei*; bai1, bai2, bai3, bai4, bai5—the first, second, third, fourth, and fifth TSP1 modules of the human brain-specific angiogenesis inhibitor 1 precursor. Residues conserved in more than 50% of the aligned sequences are highlighted by a black background, chemically similar residues present in more than 50% of the sequences are shown with a gray background. Chemically similar residues are grouped as follows: F, Y, W; I, L, V, M; R, K; D, E; N, Q; G, A; T, S.

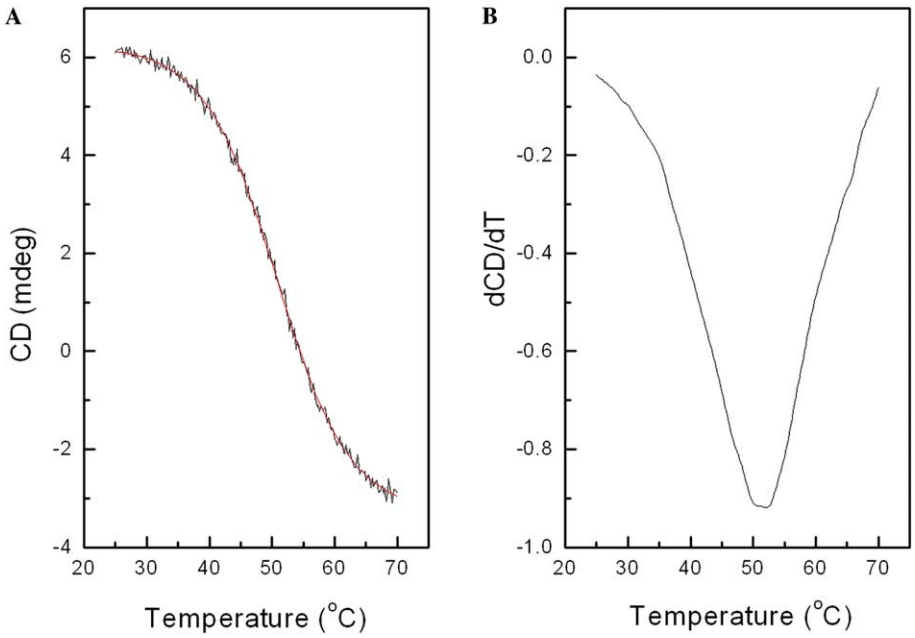

Fig. 3. Temperature dependence of the CD spectra of recombinant TSP1 domain of human thrombospondin 1. Changes in the CD of the protein were monitored at 229 nm in 10 mM Tris–HCl buffer, pH 8.0, during the course of heating from 20 to 90 °C at a heating rate of 50 °C h^{-1} (A). Melting temperature was determined by derivative processing using the J-700 program for Windows Standard Analysis Ver 1.30.00.7.JASCO (B).

Fragment B contained peptide Ser15–Leu31 and Lys52–Ser62, connected by two pairs of disulfide bonds. Amino acid sequencing of this fragment yielded PTH-cystine peaks in the third and eighth Edman degradation cycles. Since Edman degradation of disulfide-bonded peptides yields a PTH-derivative from a half-cystine residue only when the other half is also released, the data are consistent only with a Cys17–Cys54 and Cys21–Cys59 disulfide-bond pattern (cf. Table 1). In summary, the six cysteines (Fig. 2) of the TSP1 module of thrombospondin 1 are connected in a 1–5, 2–6, 3–4 disulfide-bond pattern. It is noteworthy that there are some members of the TSP1 module-family which lack certain pairs of cysteines that form disulfide-bonds. For example, the TSP1 domain of csp_plaba (Fig. 2) lacks the cysteines forming disulfide bond 3–4, whereas the cysteines forming the 2–6 disulfide bond are missing from the C-terminal TSP1 modules of complement proteins C6, C7, C8α, and C8β (Fig. 2).

Recently, it has been suggested that HB-GAM and other members of the pleiotrophin/midkine family consist of two tandem TSP1-related domains [9]. However, the two domains of pleiotrophins/midkines share only very low sequence similarity with typical TSP1 domains: the scores are less significant than the threshold values required by the SMART ([3]; http://smart.embl-heidelberg.de/) or Pfam ([22], 2002, http://www.sanger.ac.uk/Software/Pfam/index.shtml) search tools for inclusion in the TSP1 module-family. If pleiotrophins are considered to be members of the TSP1-module family, their three-dimensional structure may have diverged significantly from that of other TSP1-modules.

NMR spectroscopy of the two tandem domains of midkine and HB-GAM have shown that each domain consists of three antiparallel β-strands and the CD spectra of HB-GAM is also consistent with the predominance of β-strands in this protein [9,10]. It is noteworthy that the CD spectra of HB-GAM differs markedly from that of the second TSP1 module of thrombospondin inasmuch as it lacks the marked maxima of positive ellipticity at 212 and 230 nm [9]. It seems thus safe to assume that the three-dimensional structure of the second TSP1 domain of thrombospondin differs significantly from the structure of the two tandem domains of the pleiotriophin/midkine family.

Table 1
Amino acid sequence of the disulfide-bonded peptides derived from the TSP1-module of human thrombospondin 1 by digestion with chymotrypsin

Fragment	Cysteines connected	Peptides linked by disulfide bond
#A	3–4	CNSPSPQMNGKPCEGEARET
#B	1–5	SSCSVTCGDGVITRIRL
	2–6	KACKKDACPRS

Acknowledgments

This work was supported by Grants OTKA T022949 and OTKA T014642.

References

[1] J. Lawler, R.O. Hynes, The structure of human thrombospondin an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins, J. Cell Biol. 103 (1986) 1635–1648.

[2] L. Patthy, Detecting distant homologies of mosaic proteins. Analysis of the sequences of thrombomodulin, thrombospondin, complement components C9, C8α, and C8β, vitronectin and plasma cell membrane glycoprotein PC-1, J. Mol. Biol. 202 (1988) 689–696.

[3] I. Letunic, L. Goodstadt, N.J. Dickens, T. Doerks, J. Schultz, R. Mott, F. Ciccarelli, R.R. Copley, C.P. Ponting, P. Bork, Recent improvements to the SMART domain-based sequence annotation resource, Nucleic Acids Res. 30 (2002) 242–244.

[4] J. Lawler, Thrombospondin-1 as an endogenous inhibitor of angiogenesis and tumor growth, J. Cell. Mol. Med. 6 (2002) 1–12.

[5] M.L. Iruela-Arispe, M. Lombardo, H.C. Krutzsch, J. Lawler, D.D. Roberts, Inhibition of angiogenesis by thrombospondin-1 is mediated by 2 independent regions within the type 1 repeats, Circulation 100 (1999) 1423–1431.

[6] K. Bein, M. Simons, Thrombospondin type 1 repeats interact with matrix metalloproteinase 2. Regulation of metalloproteinase activity, J. Biol. Chem. 275 (2000) 32167–32173.

[7] H. Nishimori, T. Shiratsuchi, T. Urano, Y. Kimura, K. Kiyono, K. Tatsumi, S. Yoshida, M. Ono, M. Kuwano, Y. Nakamura, T. Tokino, A novel brain-specific p53-target gene, BAI1, containing thrombospondin type 1 repeats inhibits experimental angiogenesis, Oncogene 15 (1997) 2145–2150.

[8] K.F. Smith, K.F. Nolan, K.B. Reid, S.J. Perkins, Neutron and X-ray scattering studies on the human complement protein properdin provide an analysis of the thrombospondin repeat, Biochemistry 30 (1991) 8000–8008.

[9] I. Kilpelainen, M. Kaksonen, H. Avikainen, M. Fath, R.J. Linhardt, E. Raulo, H. Rauvala, Heparin-binding growth-associated molecule contains two heparin-binding β-sheet domains that are homologous to the thrombospondin type 1 repeat, J. Biol. Chem. 275 (2000) 13564–13570.

[10] W. Iwasaki, K. Nagata, H. Hatanaka, T. Inui, T. Kimura, T. Muramatsu, K. Yoshida, M. Tasumi, F. Inagaki, Solution structure of midkine, a new heparin-binding growth factor, EMBO J. 16 (1997) 6936–6946.

[11] T. Lukacsovich, I. Boros, P. Venetianer, New regulatory features of the promoters of an *Escherichia coli* rRNA gene, J. Bacteriol. 169 (1987) 272–277.

[12] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.

[13] H. Mach, C.R. Middaugh, R.V. Lewis, Statistical determination of the average values of the extinction coefficients of tryptophan and tyrosine in native proteins, Anal. Biochem. 200 (1992) 74–80.

[14] N. Sreerama, S.Y. Venyaminov, R.W. Woody, Estimation of the number of α-helical and β-strand segments in proteins using CD spectroscopy, Protein Sci. 8 (1997) 370–380.

[15] N. Sreerama, R.W. Woody, Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set, Anal Biochem. 287 (2000) 252–260.

[16] N. Sreerama, S.Y. Venyaminov, R.W. Woody, Analysis of protein circular dichroism spectra based on the tertiary structure classification, Anal Biochem. 299 (2001) 271–274.

[17] B. Rost, C. Sander, Combining evolutionary information and neural networks to predict secondary structure, Proteins 19 (1994) 55–77.

[18] B. Rost, in: C. Rawlings, D. Clark, R. Altman, L. Hunter, T. Lengauer, S. Wodak (Eds.), The Third International Conference on Intelligent Systems for Molecular Biology, AAAI Press, Menlo Park, CA, 1995, pp. 314–321.

[19] B. Rost, PHD: predicting one-dimensional protein structure by profile-based neural networks, Meth. Enzymol. 266 (1996) 525–539.

[20] C.A. Smith, M.K. Pangburn, C.W. Vogel, H.J. Muller-Eberhard, Molecular architecture of human properdin, a positive regulator of the alternative pathway of complement, J. Biol. Chem. 259 (1984) 4582–4583.

[21] M.V. Perdikoulis, U. Kishore, K.B. Reid, Expression and characterisation of the thrombospondin type 1 repeats of human properdin, Biochim. Biophys. Acta 1548 (2001) 265–277.

[22] A. Bateman, E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S.R. Eddy, S. Griffiths-Jones, K.L. Howe, M. Marshall, E.L. Sonnhammer, The pfam protein families database, Nucleic Acids Res. 30 (2002) 276–280.