Received September 30, 1994

Pages 497-505

ISOLATION AND CHARACTERIZATION OF A cDNA THAT CODES FOR A LIM-CONTAINING PROTEIN WHICH IS DEVELOPMENTALLY REGULATED IN HEART*

Stephen K.W. Tsui, Nathan Y.H. Yam, Cheuk Yu Lee and Mary M.Y. Waye¹

Department of Biochemistry, The Chinese University of Hong Kong, Hong Kong

SUMMARY: During our human heart cDNA sequencing project, we have obta	ain

SUMMARY: During our human heart cDNA sequencing project, we have obtained a novel cDNA clone which is very similar in DNA and amino acid sequences to a rat/mouse cysteine-rich intestinal protein (1). Sequence analysis has shown that this human cysteine-rich heart protein (hCRHP) is a protein of 77 amino acids and possesses a LIM motif which is considered to be able to bind zinc. Northern blot analyses have shown that its mRNA level in rat heart is regulated developmentally. We have expressed hCRHP in E. coli using pAED4 as the vector and the cDNA was engineered so that the authentic protein is produced. The protein was partially purified and was shown to be a basic protein.

© 1994 Academic Press, Inc.

Zinc finger proteins can be classified into at least 2 different classes based on the consensus sequences (2). One of the classes of zinc finger proteins bears a LIM motif which is constituted from a C₂HC and a C₄ motif (3-4). LIM is an abbreviated form derived from the names of three genes which share a consensus cysteine-rich sequence: lin-11 (3), Isl-1 (4) and mec-3 (5). Each LIM motif can accommodate two zinc atoms (6) and it is thought that some proteins bearing LIM motif may be involved in zinc transport (7) and/or DNA binding (3-4). Usually, LIM-containing proteins carry two LIM motifs and a homeodomain (8). However, a protein isolated from the small intestine of rat/mouse called cysteine-rich intestinal protein (CRIP) has only one LIM motif and has no homeodomain (1). Apart from the small intestine of rat/mouse, CRIP can also be found in other tissues, such as lung, adrenal, testis and spleen (1). DNA sequences similar to the CRIP gene may be existent in many different organisms (1).

We have initiated a human heart cDNA sequencing project which is aimed at revealing the expression profile of human heart and to identify novel, previously uncharacterized genes active in

Abbreviations: cysteine-rich protein, CRP; cysteine-rich intestine protein, CRIP; human cysteine-rich heart protein, hCRHP; human cysteine-rich protein, hCRP; chicken cysteine-rich protein, cCRP; 17β-estradiol-stimulated protein, ESP-1.

^{*}Accession No. of hCRHP: GenBank U09770.

¹To whom correspondence should be addressed. Address: Room 302C, Department of Biochemistry, Basic Medical Science Building, The Chinese University of Hong Kong, Shatin, Hong Kong. Fax: (852)6035123.

the human heart (9-11). A number of cDNA clones identified matched with some non-human DNA sequences. One of the these has DNA sequence resembling rat/mouse CRIP. We report here the cloning, sequencing, tissue distribution, change in expression level of hCRHP in heart during development and the expression of hCRHP in *E. coli*.

MATERIALS AND METHODS

DNA sequencing of hCRHP cDNA clone. Partial sequencing of cDNA clones isolated from a human adult heart cDNA library (λgt11 vector, Clontech) was conducted as described previously (10-11). Briefly, A bacteriophage cDNA clone was first amplified by PCR using λprimers flanking the EcoRI site of the vector (Forward: 5'ATTGGTGGCGACGACTCC TGGA3'; Reverse: 5'TTTGACACCAGACCAACTGGTA3'). The PCR product was sequenced using a cycle sequencing kit purchased from Stratagene. The sequencing products were run in a Pharmacia ALF DNA sequencer. Sequence comparisons against the GenBank and EMBL nucleotide and protein databases were performed using the BLAST electronic mail server (12). The complete sequence of a cDNA clone named human cysteine-rich heart protein (hCRHP) which matches with the mouse CRIP was then determined by cycle sequencing and was verified by manual sequencing using dideoxy chain termination method (13). The DNA and the predicted amino acid sequences were analysed by MacDNASIS from Hitachi.

Northern hybridization. Total RNA was isolated from rat tissues using guanidine thiocyanate coupled with cesium chloride centrifugation (17). About 30µg total RNA from each tissue was resolved in 1.5% agarose/2.2M formaldehyde gel. The RNAs were then transferred onto Nylon membranes (Amersham) and fixed by baking under vacuum at 80°C for 2 hours. A pair of primers within the coding region of the hCRHP cDNA was designed to give a PCR product of 204 base pairs. Radioactive random-primed probe was made using the purified PCR product as template. Northern hybridization was done by hybridizing the membrane with the [32P]-labelled probe at 42°C in the presence of 50% formamide for 20 hours. The membrane was washed to remove nonspecific annealing. Autoradiography was performed at -70°C. Densitometry of the autoradiographs and peak area integration provided relative quantification of the amount of bound cDNA probe.

RESULTS AND DISCUSSION

The DNA and protein sequences of hCRHP. Among the various non-human match clones that we sequenced, a cDNA clone which matches with a rat/mouse CRIP was identified. It was named the human cysteine-rich heart protein (hCRHP) (Fig.1). Excluding the vector sequence and the poly A region, the cDNA insert is 416 base pairs in length. The initiation codon was found at nucleotide number 65. The stop codon is at the nucleotide number 296. The consensus initiation sequence CCRCCAUG (R represents purine) (18) is present in the start of the open reading frame (ORF). By aligning the DNA sequences of the ORF of rat/mouse CRIP and hCRHP, they were found to have a similarity of 88.7% (Fig. 2). We believe that rat/mouse CRIP and hCRHP are

1	A	GAG	TCT	CGC	ACT	GTA	GCC	CGT	GCC	GCC	CCA	GCC	GCT	GCC	GCC	TGC	46
1 47	ACC	GGA	ccc	GGA	GCC	GCC	M ATG	P CCC	K AAG	TGT	P CCC	K AAG	C TGC	N AAC	K AAG	E GAG	10 94
	V GTG																26 142
27 143	CCC	C TGC	L CTG	K AAG	r c c	E GAG	K AAA	c tgt	G GGG	X AAG	T ACG	L CTG	T ACC	S TCT	G GGG	G GGC	42 190
			_			_		_					-	_			
	H CAC																58 238
191 59		GCT M	gag F	CAC G	gaa P	GGC K	AAA G	CCC F	TAC G	TĞC R	AAC G	CAC G	CCC	TGC E	TAC S	GTA H	
191 59 239 75	CAC A	GCT M ATG F	GAG F TTT K	G GGG *	GAA P CCT	GGC K AAA	G GGC	CCC F TTT	TAC G GGG	TĞC R CGG	AAC G GGC	G G GGA	CCC A GCC	tēc e gag	TAC S AGC	GTA H CAC	238 74
191 59 239 75 287	CAC A GCC T	GCT M ATG F TTC	GAG F TTT K AAG	GGG *	GAA P CCT ACC	GGC K AAA AGG	G GGC TGG	F TTT TGG	G GGG AGA	TĞC R CGG CCC	G GGC ATC	G GGA CTT	CCC A GCC GGC	TGC E GAG TGC	S AGC TTG	GTA H CAC CAG	238 74 286 78 334

Fig. 1. The cDNA and amino acid sequences of hCRHP. In the amino acid sequence, the amino acid residues underlined (CCHCCCC) constitute the LIM motif. In the DNA sequence, the start (ATG) and stop (TAA) codons are underlined. The poly A tail at the end is not shown.

structurally and functionally related because they share a high degree of similarity in their ORF. Unlike the ORF region, when the 5' untranslated region (5'UTR) and the 3' untranslated region (3'UTR) of the two genes were aligned, only 52% (32 out of 61 bps) and 48% (54 out of 112 bps) homology were obtained respectively.

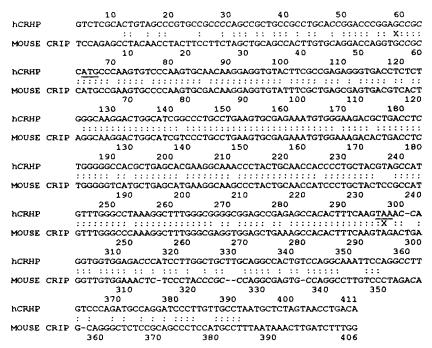


Fig. 2. Alignment of DNA sequences of CRIP and hCRHP. The alignment was performed using MacDNASIS. Nucleotides that are identical between the two sequences are marked by ':' while those that are different are left blank.

	10	20	30	40	50	60
hcrhp	MPKCPKCNKEVYFAER\	TSLGKO	WHRPCLKCEKCO	KTLTSGGHA	EHEGKPYCNH	PCYVAM
	X::::::::::::::	::::::	:::::::::::::	::::::::	:::::::::	::: ::
MOUSE CRIP	MPKCPKCDKEVYFAER\	TSLGKD	WHRPCLKCEKCO	KTLTSGGHA	EHEGKPYCNH	PCYSAM
	10 70	20	30	40	50	60
hCRHP	FGPKGFGRGGAESHTF	ς				
	:::::::::::::::::::::::::::::::::::::::	ζ				
MOUSE CRIP	FGPKGFGRGGAESHTF1	ζ.				

Fig. 3. Alignment of amino acid sequences of CRIP and hCRHP. The alignment was performed using MacDNASIS. Amino acids that are identical between the two sequences are marked by '.' while those that are similar were marked by '.'. Those amino acids that are different between the two sequences are left blank.

After translating the ORF of the hCRHP cDNA clone, a protein sequence of 77 amino acids was obtained. When the amino acid sequences of CRIP and hCRHP were aligned, 97.4% identity between these two proteins was found (Fig. 3), with only a difference of two amino acids. At position 8, asparagine (N) in hCRHP was replaced by aspartic acid (D). At position 58, valine (V) in hCRHP was replaced by serine (S). The two proteins have an even higher homology than that of the DNA. The calculated molecular weight is 8561 daltons. It possesses a LIM motif which is constituted by 7 cysteines and 1 histidine. The first finger is at position cysteine 4 to cysteine 28 and the second one is at position cysteine 31 to cysteine 56. The hydrophobic residues phenylalanine 13 and leucine 20 are invariant in most zinc fingers having two cysteine residues and two histidine residues, such as TFIIIA (finger one to three) and Xfin 31 (19). hCRHP has an excess of basic residues (10 lysyl, 5 histidyl and 3 arginyl) over acidic ones (1 aspartyl and 6 glutamyl). Therefore the estimated pI as determined by the software MacDNASIS (Hitachi) was 8.83. The amino acid sequences of the two fingers is 75.5% identical (40 out of 53) and 83.0% similar (44 out of 53) to the rat 17β-estradiol-stimulated protein (ESP1) gene (20). Another striking similarity between these two proteins is that administration of dexamethasone to neonatal rat caused the rise of CRIP mRNA content (21) while estradiol can increase the mRNA level of ESP1 in brain of adult rat (20).

A glycine-rich domain in human cysteine-rich protein (hCRP) and rat/mouse CRIP has been reported (22). We have observed a special feature of the glycine-rich domain of hCRHP. There are 7 hydrophobic residues, 5 glycine residues and 2 basic residues from valine 58 to alanine 71. The arrangement is HHHHGH+GHG+GGH (H:hydrophobic; G:glycine; +:basic residue). The two basic residues are located within the five glycine residues which are in turn surrounded by hydrophobic residues.

A human protein hCRP has been characterized (22) and it has been proposed that the adjacent finger in the finger doublet motif is a result of the duplication of the finger doublet region of the rat/mouse CRIP gene. Also, rCRP2 was identified to be the rat homolog of hCRP (23). Our report suggests that hCRHP is the homolog of rat/mouse CRIP and that hCRHP and hCRP are distinct genes even though they may share a common evolutionary history (Fig. 4). The percentage homology of the LIM motifs among various proteins is shown in Table 1.

Northern hybridization of hCRHP. It was found that hCRHP mRNA is expressed at a high level in human fetal heart. By comparing the location of the band with that of 18S and 28S

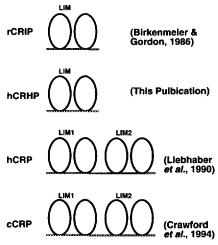


Fig. 4. Schematic representation of CRIP, hCRP, hCRP and cCRP. The oval regions represent the LIM motif of the proteins. (For details of each protein, please see the references. Crawford *et al.*, 1994 in reference no. 28.)

ribosomal RNA, we found that the mRNA has a size of about 670 nucleotides (Fig. 5A). It is of a size similar to the hCRHP cDNA sequence we have obtained. Since the DNA sequences between rat/mouse CRIP and hCRHP were very similar, hCRHP cDNA probe was hybridized with total RNA of various rat tissues (Fig. 5B). It was shown that the small intestine has the highest signal. Lower signals could be detected in adult heart and spleen. Virtually no signal could be detected in skeletal muscle, kidney and liver. The results obtained agree with those previously published (21). When rat neonatal (one day old) and adult heart RNAs were hybridized with the hCRHP cDNA probe, only a weak signal for rat adult heart was observed (Fig. 5C); however, we found that the mRNA level of hCRHP in neonatal heart is 5 times more than that of adult heart from densitometry. Staining of the membrane with methylene blue shows that virtually equal amounts of total RNA were loaded in the different lanes (data not shown). On the other hand, rat/mouse CRIP mRNA level in small intestine of adult rat was six times more than that of one day old neonatal rat (1). Our result suggests that the expression of hCRHP changes during development.

Table 1. Comparison of similarity of LIM motifs in various proteins. The LIM motif of proteins in the left column were compared with that in the right row. The denominator represents the number of amino acid residues compared while the nominator represents the number of amino acid residues of the proteins in the right row that are identical to the amino acid residues of the proteins in the left column.

	rCRIP	hCRHP	hCRP(LIM1)	hCRP(LIM2)	cCRP(LIM1)	cCRP(LIM2)
rCRIP	-	49/51	21/51	24/51	21/51	24/51
hCRHP	-	-	21/51	25/51	22/51	25/51
hCRP(LIM1)	-	-	-	27/50	47/50	27/50
hCRP(LIM2)	_	-	-		28/50	46/50
cCRP(LIM1)	-	-	-	-	-	28/50
cCRP(LIM2)	-	-	-	-	-	-

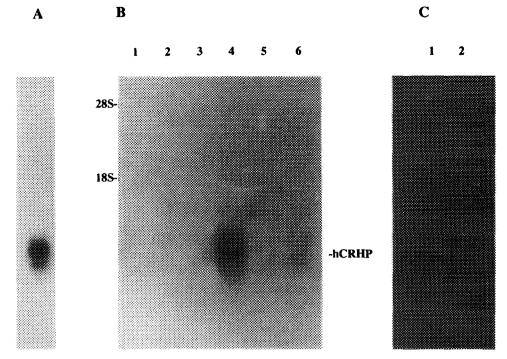


Fig. 5. Tissue distribution of hCRHP in human neonatal heart and in rat tissues. (A). Northern hybridization of hCRHP probe to human neonatal heart total RNA. (B). Northern hybridization of hCRHP probe to adult rat RNAs. Lane 1: rat heart. Lane 2: rat skeletal muscle. Lane 3: rat liver. Lane 4: rat small intestine. Lane 5: rat kidney. Lane 6: rat spleen. (C). Northern hybridization of neonatal and adult rat heart total RNA. Lane 1 is adult heart total RNA. Lane 2 is neonatal heart total RNA.

Expression of hCRHP. hCRHP cDNA clone was successfully amplified using a tailor-made cloning primer and an oligo dT primer (Fig. 6). Thus, the oligo dT primer can be used as a common 3' cloning primer for amplifying and cloning interesting cDNA clones. This strategy is especially suitable for a large cDNA sequencing and expression project. Since the oligo dT primer can prime along the poly A region of the 3' end of the template, a set of PCR products of a range of sizes were made. However, this will not introduce any problem for later experiments because, after being subcloned into pAED4, a single hCRHP cDNA clone can be isolated easily.

The success of the directional cloning was proven by restriction cutting of the putative recombinant plasmid, PCR using primers complementary to the internal sequence of hCRHP cDNA and manual sequencing. When we attempt to transform the recombinant plasmid (pAED4-hCRHP) into *E. coli.* BL21(DE3) strain, no transformed colony was obtained (data not shown). It was suggested that hCRHP produced inside the bacteria chelated metal ions such as zinc and copper and prevented the normal growth of bacteria. Thus, a high level of hCRHP was very toxic to bacteria. In order to circumvent this problem, *E. coli* BL21(DE3)pLysE and BL21(DE3)pLysS strains were used which guaranteed minimal production of hCRHP under uninduced conditions (14). However, out of 16 randomly picked successful transformants, only one produced a minute amount of cloned protein upon induction (data not shown). To overcome this problem, 0.2mM

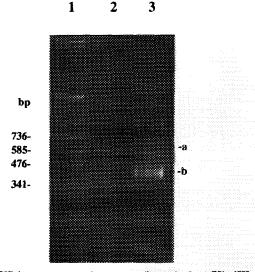


Fig. 6. Subcloning of hCRHP into an expression vector. Lane 1: the λ/HindIII-pUC18/Sau3AI size marker. Lane 2: the PCR product of hCRHP using λ PCR primers. Lane 3: the PCR products of hCRHP using specific cloning primer and an oligo dT primer. (a) and (b) represent the location of PCR products containing hCRHP.

ZnSO₄ was supplemented during induction (15), and 8 out of 8 randomly picked successful transformants produced large amounts of the cloned protein upon induction. A representative clone is shown in Figure 6A. It could be inferred from our findings that zinc was required for the successful production of hCRHP upon induction.

After resolving the crude bacterial extract in 15% SDS-PAGE, the proteins were transferred onto a nylon membrane and stained with Coomassie Blue (Fig. 7A). It can be shown that the intense band at about 8.5 kDa (lane 3, Fig. 7A) was the protein coded by hCRHP cDNA and was absent when not induced by IPTG. Its identity has also been verified by sequencing 15 amino acid residues from the N-terminal of the protein (data not shown). When the crude bacterial extract was extracted once with 0.25M HCl (24), hCRHP was mainly found in the acid soluble fraction (Fig. 7B). It was shown that hCRHP is a basic protein and has a solubility property similar to other histones and acid soluble nuclear proteins (24). By acid extraction, hCRHP can be partially purified. It can serve as a first step for purifying hCRHP.

The possible roles of CRIP and hCRHP. CRIP was thought to be a zinc carrier which can bind to zinc during transmucosal zinc transport of rats (7). Some results suggested that CRIP and intestinal metallothionein competitively bind zinc during zinc transport (25). Alternatively, the observation of zinc transport may be explained if CRIP has the properties of metal exchange (15). Also, it has been shown that CRIP mRNA level is not primarily dependent on zinc level (26). On the other hand, the administration of dexamethasone, a glucocorticoid, would lead to the precocious rise of CRIP mRNA level in rat neonate (21). Together with the information that the level of CRIP mRNA in the small intestine of rat changes with different developmental stages of

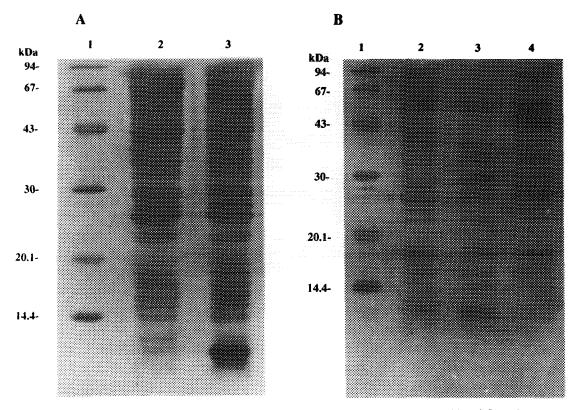


Fig. 7. Expression of hCRHP. (A). SDS-PAGE of hCRHP with 15% polyacrylamide gel. Lane 1 is molecular size marker. Lane 2: uninduced recombinant bacterial crude extract. Lane 3: induced recombinant bacterial crude extract. (B). Acid extraction of hCRHP. Lane 1: the molecular size marker. Lane 2: the uninduced recombinant bacterial crude extract. Lane 3: the acid soluble fraction of induced recombinant bacterial crude extract. Lane 4: the acid insoluble fraction of induced recombinant bacterial crude extract.

small intestine (1), it was postulated that CRIP may mediate the action of glucocorticoid in the development of the small intestine (27). At this moment, the exact function of CRIP is still under intense investigation. Our results show that hCRHP, the human homolog of rat/mouse CRIP, is developmentally regulated in rat heart in a pattern different from that of rat CRIP in the rat small intestine. This result supports the idea that CRIP or hCRHP is not a intestinal-specific protein. The change in expression during development, excess of basic residues and existence of zinc finger motifs suggect that hCRHP may be involved in transcriptional regulation of gene expression during the development of heart.

ACKNOWLEDGMENTS

The work was supported by a UPGC grant RGC Ref. No.: CUHK16/93M. We would like to thank Don Doering and Paul Matsudaira for giving us the pAED4 expression vector, and Dr. Louise Chow for providing us the tissues.

REFERENCES

- 1. Birkenmeier, E.H. and Gordon, J.I. (1986) Proc. Natl. Acad. Sci. USA 88, 2516-2520.
- 2. Evans, R.M. and Hollenberg, S.M. (1988) Cell 52, 1-3.
- 3. Freyd, G., Kim, S.K. and Horvitz, H.R. (1990) Nature 344, 876-879.
- Karlsson, O., Thor, S., Norberg, T., Ohlsson, H. and Edlund, T. (1990) Nature 344, 879-882.
- 5. Way, J.C. and Chalfie, M. (1988) Cell 54, 5-16.
- Michelsen, J.W., Schmeichel, K.L., Beckerle, M.C. and Winge, D.R. (1993) Proc. Natl. Acad. Sci. USA 90, 4404-4408.
- 7. Hempe, J.M. and Cousins, R.J. (1991) *Proc. Natl. Acad. Sci. USA* 88, 9671-9674.
- 8. Li, P.M., Reichert, J., Freyd, G., Horvitz, H.R. and Walsh, C.T. (1991) *Proc. Natl. Acad. Sci. USA* 88, 9210-9213.
- 9. Liew, C.C., Hwang, D.M., Fung, Y.W., Laurenssen, C., Cukerman, E. and Lee, C.Y. (1994) *Proc. Natl. Acad. Sci. USA* in press.
- 10. Liew, C.C. (1993) J. Mol. Cell. Cardiol. 25, 891-894.
- 11. Tsui, K.W., Fung, Y.W., Lam, W.Y., Waye, M.M.Y., Lee, C.Y. and Liew, C.C. (1994) *J. Mol. Cell. Cardiol.* 26, CLXXVI.
- 12. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403-410.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- 14. Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Methods in Enzymology 185, 60-89.
- 15. Kosa, J.L., Michelsen, J.W., Louis, H.A., Olsen, J.I., Davis, D.R., Beckerle, M.C. and Winge, D.R. (1994) *Biochemistry* 33, 468-477.
- 16. Laemmli, U.K. (1970) Nature 227, 680-685.
- 17. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual 2nd. ed. Cold Spring Harbor Labortory Press, Cold Spring Harbor, New York.
- 18. Kozak, M. (1983) Microbiol. Rev. 47, 1-45.
- 19. Pabo, C.O. (1992) Annu. Rev. Biochem. 61, 1053-1095.
- Nalik, P., Panayotova-Heiermann, M. and Pongs, O. (1989) Mol. Cell. Endocrinol. 62, 235-242.
- Levenson, C.W., Shay, N.F., Lee-Ambrose, L.M. and Cousins, R.J. (1993) Proc. Natl. Acad. Sci. USA 1993, 712-715.
- 22. Liebhaber, S.A., Emery, J.G., Urbanek, M., Wang, X. and Cooke, N.E. (1990) *Nucl. Acids Res.* 18, 3871-3879.
- Okano, I., Yamamoto, T., Kaij, A., Kimura, T., Mizuno, K. and Nakamura, T. (1993) FEBS Lett. 333, 51-55.
- 24. Liew, C.C. and Cukerman, E. (1993) Mol. Cell. Biochem. 121, 175-179.
- 25. Hempe, J.M. and Cousins, R.J. (1992) J. Nutr. 122, 89-95.
- Levenson, C.W., Shay, N.F., Hempe, J.M. and Cousins, R.J. (1994) J. Nutr. 124, 13-17.
- Needleman, D.S., Leeper, L.L., Nanthakumar, N.N. and Henning, S.J. (1992) J. Pediatr. Gastroenterol. Nutr. 16, 15-22.
- 28. Crawford, A.W., Pino, J.D. and Beckerle, M.C. (1994) J. Cell Biol. 124, 117-127.