Cloning and sequence analysis of rat bone sialoprotein (osteopontin) cDNA reveals an Arg-Gly-Asp cell-binding sequence

(cell adhesion/bone matrix proteins/glycosylation)

ÅKE OLDBERG, AHNDERS FRANZÉN, AND DICK HEINEGÅRD

Department of Physiological Chemistry, University of Lund, P.O. Box 94, S-221 00 Lund, Sweden

Communicated by Peter Reichard, May 28, 1986

ABSTRACT The primary structure of a bone-specific sialoprotein was deduced from cloned cDNA. One of the cDNA clones isolated from a rat osteosarcoma (ROS 17/2.8) phage λ gt11 library had a 1473-base-pair-long insert that encoded a protein with 317 amino acid residues. This cDNA clone appears to represent the complete coding region of sialoprotein mRNA, including a putative AUG initiation codon and a signal peptide sequence. The amino acid sequence deduced from the cDNA contains several Ser-Xaa-Glu sequences, possibly representing attachment points for O-glycosidically linked oligosaccharides and one Asn-Xaa-Ser sequence representing a likely site for the N-glycosidically linked oligosaccharide. An interesting observation is the Gly-Arg-Gly-Asp-Ser sequence, which is identical to the cell-binding sequence identified in fibronectin. The presence of this sequence prompted us to investigate the cell-binding properties of sialoprotein. The ROS 17/2.8 cells attached and attained a spread morphology on surfaces coated with sialoprotein. We could demonstrate that synthetic Arg-Gly-Asp-containing peptides efficiently inhibited the attachment of cells to sialoprotein-coated substrates. The results show that the Arg-Gly-Asp sequence also confers cell-binding properties on bone-specific sialoprotein. To better reflect the potential function of bone sialoprotein-we propose the name "osteopontin" for this protein.

Bone contains an extensive mineralized matrix of hydroxyapatite crystals. The major organic constituent is type I collagen, forming fibers possibly important for the organization of the mineral. The tissue also contains a number of proteins including phosphoproteins (1) and osteonectin (2); other proteins appear to represent unique expressions of the osteogenic phenotype. One such protein is osteocalcin, a low molecular mass protein containing y-carboxyglutamic acid residues (3). Early, Herring and Kent isolated a sialic acid-rich protein with a molecular mass of about 25 kDa from bone and termed this sialoprotein (4). More recent work using improved techniques (5, 6) for the isolation of bone macromolecules has shown that the molecule has a molecular mass of 57 kDa and is only present in extracts of the calcified matrix of bone as measured by ELISA (6). It is synthesized in explant cultures of rat calvaria (7), probably by the osteoblasts. The sialoprotein can be separated to yield two somewhat different molecules. They differ slightly in amino acid composition. However, their peptide patterns show major similarities, although there are peptides present only in one or the other. A major difference is in the content of sialic acid and phosphate, where the smaller sialoprotein only contains some 4-5% sialic acid compared with 13-14% in the larger (6). Some of the differences noted in peptide patterns may actually be due to different substitutions of the protein. The core protein of the major sialoprotein should have an

approximate molecular mass of about 30 kDa, as calculated from the protein content and molecular weight of the whole molecule (6).

An interesting question that arose during this study was whether the sialoprotein could act as a cell-binding protein. Quite a number of such proteins present in a number of different connective tissues have been identified. Thus, cultured cells are known to adhere to, for instance, fibronectin (8), collagen (9), laminin (10), and vitronectin (11). The interaction between fibronectin and a cell-surface receptor is well-characterized and can be ascribed to the sequence Arg-Gly-Asp (12) in the cell-attachment domain of the protein (13). An identical sequence present in the plasma protein vitronectin enables this protein to promote attachment of cells (14). This sequence is also found in the α chain of fibrinogen (15) and has been shown to be essential for platelet aggregation and binding of platelets to fibrinogen (16).

The cloning and sequence determination of rat-bone sialoprotein, presented in this report, informed us about the presence of an Arg-Gly-Asp sequence. We could verify that this sequence confers the bone protein cell-binding properties. In addition, the sialoprotein binds strongly to hydroxylapatite. Therefore, we suggest that the protein is named osteopontin, denoting that it is a product of cells in the osteoid matrix and that it can form a bridge (latin pons) between cells and the mineral in the matrix.

MATERIAL AND METHODS

The rat osteosarcoma cell line ROS 17/2.8 (17) and primary cultures of rat peritoneal fibroblast (18) were cultured in Ham's F-12 medium (GIBCO) supplemented with 10% fetal bovine serum. Synthetic peptides corresponding to the cell-binding site in fibronectin (Arg-Gly-Asp) and variants thereof (12) were gifts from M. Pierschbacher and E. Ruoslahti (Cancer Research Center, La Jolla Cancer Research Foundation, La Jolla, CA). Rat-bone sialoprotein was purified, and antibodies were raised in rabbits as described (6, 7). Enzymes and materials were obtained from Pharmacia except when otherwise stated.

Construction and Screening of the cDNA Library. Total cellular RNA was extracted from the ROS 17/2.8 cells by the guanidine isothiocyanate procedure, and poly(A)⁺ RNA was isolated by chromatography on oligo(dT)-cellulose as described (19). Double-stranded cDNA was synthesized from poly(A)⁺ RNA as described by Gubler and Hoffman (20) and was ligated to phage λ gt11 arms as described by Schwarzbauer *et al.* (21). Packaging extracts were prepared, and recombinant DNA was packaged as described (22). Recombinant phage were amplified in *Escherichia coli* strain Y1088 and used to infect *E. coli* strain Y1090 (23). About 250,000 recombinants were obtained from 140 ng of double-stranded cDNA. Recombinant plaques were screened with the anti-

Abbreviations: bp, base pair(s); kb, kilobase(s).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

serum (diluted 1:100) and 125 I-labeled protein A (10⁵ cpm/ml) (14). Screening of the library with cDNA probes was performed as described (24).

Transfer Blot and Nucleotide Sequence Analysis. RNA transfer blots were done as described (19) after electrophoresis of $poly(A)^+$ RNA on 1.2% agarose gels containing formaldehyde (22). Sequence analysis was performed as described (24), and data analysis was performed with the Staden DB system sequence-analysis program (25).

Cell Attachment Assay. Microtiter polystyrene plates (Nunc, Immunoplate IF) were coated overnight at room temperature with 5 μ g of rat osteopontin per ml of 4 M guanidine HCl/10 mM Tris HCl, pH 7.4. The wells were washed with 0.15 M NaCl/5 mM sodium phosphate, pH 7.4 (phosphate-buffered saline) before addition of 100 μ l of F-12 medium containing bovine serum albumin (1 mg/ml) and dilutions of peptides. The rat osteosarcoma cells were harvested from the culture plates with 1 mM EDTA in phosphate-buffered saline, washed once in F-12 medium, and then resuspended at a concentration of 3×10^5 cells per ml in F-12 medium containing bovine serum albumin (1 mg/ml). Each well received 100 μ l of the cell suspension, and the plates were incubated for 1 hr at 37°C. Bound cells were quantitated by measuring endogenous N-acetyl- β -D-hexosaminidase (26) after cells not attached had been removed by rinsing with phosphate-buffered saline.

Amino Acid Sequence Analysis. Cyanogen bromide cleavage of osteopontin was performed in 1% (wt/vol) cyanogen bromide dissolved in 70% (wt/vol) formic acid as described (27). Three peptides were separated by chromatography on Mono-Q (Pharmacia) eluted with a 0–1 M linear NaCl gradient in 7 M urea/20 mM Tris HCl, pH 8 (data not shown). The three peptides were assayed for cell-binding properties, and the one that was active was sequenced by J. Stenflo (Department of Clinical Chemistry, General Hospital, Malmö, Sweden) with the use of a gas-phase sequencer (28).

RESULTS

Osteopontin is most likely produced by the osteoblasts. To demonstrate that the osteosarcoma cells (ROS 17/2.8) synthesize osteopontin, a set of experiments to isolate the protein from cell cultures was initiated. The cells have been identified as osteoblasts expressing receptors for parathyroid hormone and 1,25-dihydroxycholecalciferol (17). We could isolate a protein from spent osteosarcoma cell culture media that was biochemically and immunologically identical to osteopontin isolated from rat-bone tissue (unpublished observation). Since these cells synthesize osteopontin, we chose to use mRNA isolated from them to construct a cDNA library in the expression vector $\lambda gt11$. The library was screened for the presence of osteopontin clones with a specific antiserum.

Identification and Isolation of Sialoprotein cDNA Clones. Initial antibody screening of 120,000 recombinant plaques identified three putative osteopontin clones, $\lambda ROP1$, $\lambda ROP2$, and λ ROP3. After several rounds of plaque purification and rescreening with the antiserum, the clones were subjected to restriction enzyme mapping and partial nucleotide-sequence analysis. The cDNA inserts from the clones $\lambda ROP1$ and λ ROP2 were both 1.1 kilobase pairs (kbp) long and gave identical fragments after treatment with the restriction enzymes Hae III and Sau3A. A partial nucleotide sequence for each cDNA insert was determined after ligation into the EcoRI site of phage M13. Sequence analysis of about 250-300 nucleotides from both ends of the λ ROP1 and λ ROP2 inserts showed identical sequences, strongly suggesting that the two clones were identical. The cDNA insert from $\lambda ROP3$ was 0.8 kbp long and had a 5' nucleotide sequence identical to the sequence of 346 nucleotides downstream from the 5' end of

 λ ROP1 and λ ROP2. The nucleotide sequence in the 3' end of λ ROP3 contained the poly(A) tail, but a pentanucleotide sequence present in λ ROP1 and λ ROP2 just upstream of the poly(A) tail was missing (see below). The 5' and 3' sequences and the *Hae* III and *Sau*3A restriction enzyme fragment patterns indicate that the three cDNA clones identified by the antiserum were derived from a mRNA encoding a single protein.

The λ ROP1 cDNA insert hybridized to a 1.4-kilobase (kb) mRNA in RNA transfer-blot analysis of mRNA isolated from ROS 17/2.8 cells (Fig. 1). No hybridization was observed to poly(A)⁺ RNA from rat peritoneal fibroblasts, consistent with and extending previous information obtained with immunoassay, showing that osteopontin is only expressed in bone (6). As the 1.1-kbp λ ROP1 cDNA insert was too short to encompass the entire mRNA, the λ gt11 cDNA library was screened by plaque hybridization. An EcoRI/Ava II fragment (nucleotides 383-596 in Fig. 2) was isolated from λ ROP1, labeled by nick-translation, and used to screen the cDNA library. Six hybridizing plaques were identified out of 40,000 plaques screened. From these clones we chose λ ROP11, having the largest insert, for further analysis. The clone contained 1.2-kbp and 250-bp EcoRI insert fragments, indicating an EcoRI site in the cDNA. Transfer-blot analysis of $ROS 17/2.8 poly(A)^+ RNA$ showed that the two *Eco*RI insert fragments hybridized to the same 1.4-kb mRNA. This demonstrates that the two cDNA fragments were derived from the same mRNA species and excludes the possibility that the two fragments were artificially joined during the construction of the library. The λ ROP11 insert shared Hae III and Sau3A restriction enzyme sites with λ ROP1, λ ROP2, and λ ROP3, indicating that also $\lambda ROP11$ was derived from the same mRNA.

Nucleotide and Deduced Amino Acid Sequence of Osteopontin cDNA. The nucleotide sequence of the λ ROP11 1.2-kbp fragment was determined after random fragmentation on both strands (Fig. 2). The sequence of the 250-bp fragment was determined after ligation into the *Eco*RI site of phage M13. To determine the orientation of the 250-bp *Eco*RI fragment in the intact cDNA, a 393-bp *Hae* III fragment (nucleotides 114–507 in Fig. 2) was ligated into the *Sma* I site of phage M13 and sequenced. This gave the proper sequence upstream of the *Eco*RI site and also confirmed the presence of a single *Eco*RI site in λ ROP11.

The λ ROP11 insert is 1473 bp long and contains an open reading frame coding for 345 amino acids (those 28 Nterminal residues not shown in Fig. 2). A possible AUG initiation codon is situated 79 nucleotides from the 5' end of the cDNA. The CAACC sequence upstream of this putative

initiation triplet resembles a suggested consensus CC_G^ACC

sequence for eucaryotic initiation sites (30). Furthermore, after the first methionine residue there is a stretch of about 20 predominantly hydrophobic and neutral amino acid residues. This is similar to the composition found in signal-peptide

AΒ

- 1078

872

FIG. 1. Transfer blot analysis of poly(A)⁺ RNA isolated from rat peritoneal fibroblasts (lane A) and ROS 17/2.8 cells (lane B). The samples containing 2 μ g of RNA each were electrophoresed on a 1.2% agarose gel containing formaldehyde, transferred to a nitrocellulose filter, and hybridized to a ³²P-labeled λ ROP1 cDNA insert. A *Hae* III digest of phage ϕ X174 DNA was used as a size marker (shown in kbp).

GCCTCAGCATCCTTGGCTTTGCAGTCTCCTGCGGCAAGCATTCTCGAGGAAGCCAGCC	79
ATGAGACTGGCAGTGGTTTGCCTTTGCCTGTTCGGCCTTGCCTCCTGTCTCCCGGTGAAAGTGGCTGAGTTTGGC MetArgLeuAlaValValCysLeuCysLeuPheGlyLeuAlaSerCysLeuProValLysValAlaGluPheGly	154 25
AGCTCAGAGGAGAAGGCGCATTACAGCAAACACTCAGATGCTGTAGCCACTTGGCTGAAGCCTGACCCATCTCAG SerSerGluGluLysAlaHisTyrSerLysHisSerAspAlaValAlaThrTrpLeuLysProAspProSerGln	224 50
	204
LysGlnAsnLeuLeuAlaProGlnAsnSerValSerSerGluGluThrAspAspPheLysGlnGluThrLeuPro	304 75
AGCAACTCCAATGAAAGCCATGACCACATGGACGATGATGACGACGACGACGACGACGGAGGACCATGCAGAGAGC SerAsnSerAşnGluSerHisAspHisMetAspAspAspAspAspAspAspAspAspGlyAspHisAlaGluSer *	379 100
GAGGATTCTGTGAACTCGGATGAATCTGACGAATCTCACCATTCCGATGAATCTGATGAGTCCTTCACTGCCAGC GluAspSerValAsnSerAspGluSerAspGluSerHisHisSerAspGluSerAspGluSerPheThrAlaSer	454 125
ACACAAGCAGACGTTTTGACTCCAATCGCCCCCACAGTCGATGTCCCTGAC <u>GGCCGAGGTGATAGC</u> TTGGCTTAC ThrGlnAlaAspValLeuThrProIleAlaProThrValAspValProAsp <mark>GlyArgGlyAspSer</mark> LeuAlaTyr	529 150
GGACTGAGGTCAAAGTCCAGGAGTTTCCCTGTTTCTGATGAACAGTATCCCGATGCCACAGATGAGGACCTCACC GlyLeuArgSerLysSerArgSerPheProValSerAspGluGlnTyrProAspAlaThrAspGluAspLeuThr	604 175
TCCCGCATGAAGAGCCAGGAGTCCGATGAGGCTATCAAGGTCATCCCAGTTGCCCAGCGTCTGAGCGTGCCCTCT SerArgMetLysSerGlnGluSerAspGluAlaIleLysValIleProValAlaGlnArgLeuSerValProSer	679 200
GATCAGGACAGCAACGGGAAGACCAGCCATGAGTCAAGTCAGCTGGATGAACCAAGCGTGGAAACACACAGCCTG AspGlnAspSerAsnGlyLysThrSerHisGluSerSerClnLeuAspGluProSerValGluThrHisSerLeu	754 225
GAGCAGTCCAAGGAGTATAAGCAGAGGGCCAGCCACGAGAGCACTGAGCAGTCGGATGGAT	829 250
AAGCCGGATGCAATCGATAGTGCAGAGCGGTCGGATGCTATCGACAGTCAGGCGA&TTCCAAAGCCAGCCTGGAA LysProAspAlalleAspSerAlaGluArgSerAspAlaİleAspSerGlnAlaSerSerLysAlaSerLeuGlu	904 275
CATCAGAGCCACGAGTTTCACAGCCATGAGGACAAGCTAGTCCTAGACCCTAAGÁGTAAGGAAGATGATAGGTAT HisGlnSerHisGluPheHisSerHisGluAspLysLeuValLeuAspProLysSerLysGluAspAspArgTyr	979 300
CTGAAATTCCGCATTTCTCATGAATTAGAGAGTTCATCTTCTGAGGTCAATTAAAGAAGAGGCAAAACCACAGTT LeuLysPheArgIleSerHisGluLeuGluSerSerSerGluValAsn	1054
CCTTACTTTGCTTTAAAAAAAAAAAAAAAAAAAAAAAAA	1129
TGGATACATGTATGTGGAGAAAGAAATAGATAGTGTTTTGGGCCCTGAGCTTAGTTCGTTGTTTCATGCAGACAC	1204
CACTGTAACCTAGAAGTTTCAGCATTTCGCTTCTGTTCTTTCT	1279
ATGATTGCTATTCTTTTATGAATAAAATGTATGTAGAGGCAGGC	354
AACTATAATAGTCTGTGTCACTATAATCTTTTGGTTTTATAATTAGTGTATATTTTGTTGTGATTATT	1429

sequences (31). Based on these observations, we assigned this methionine as the first residue in the predicted osteopontin sequence. If one assumes that the signal peptide is cleaved as suggested in Fig. 2, the molecular mass of osteopontin would be 32,627 daltons. This agrees well with the previously calculated molecular mass of \approx 30 kDa (6). The amino acid composition deduced from the cDNA sequence analysis is in good agreement with the amino acid composition determined for rat osteopontin (Table 1). The amino acids serine, glutamic acid/glutamine, and aspartic acid/aspargine together constitute approximately half of the amino acids in the protein. This characteristic and unusual amino acid composition, found both by direct analysis (7) and deduced from the nucleotide sequence, is strong evidence for the identity of the clone as being derived from osteopontin mRNA.

Contributing to the characteristic amino acid composition is a highly acidic sequence consisting of nine consecutive aspartic acid residues (amino acids 90-98 in Fig. 2). A

amino acid sequence of rat osteopontin. The nucleotide sequence of $\lambda ROP11$ was determined. The numbering of amino acids starts with a methionine residue encoded by a putative ATG initiation codon. A possible signal-peptide cleavage site (29) is indicated by an arrow. The cell-attachment sequence is boxed. The segment containing three head-to-tail repeated sequences is underlined, and the possible site of Nlinked glycosylation is indicated by a star. Four potential polyadenylylation signal sequences are shown overlined. The pentanucleotide sequences between the arrow heads was absent in the sequence of λ ROP11 and was found only in the clones $\lambda ROP1$ and $\lambda ROP2$. The sequence underlined with a dashed line was confirmed by amino acid sequence analysis.

FIG. 2. Nucleotide and deduced

Gly-Arg-Gly-Asp-Ser sequence was deduced from the nucleotide sequence (boxed in Fig. 2). This pentapeptide sequence is identical to the cell-attachment sequence in fibronectin (14) and contains the Arg-Gly-Asp sequence, which has been shown to be essential for the cell adhesion-promoting activity of several proteins as discussed above. Apart from this pentapeptide, no sequence homologies with other proteins were found in the data base of Barker et al. (32).

Analysis of internal repeats by the DIAGON program (33) revealed a sequence nine amino acid residues long that is repeated three times in the protein (underlined in Fig. 2). This head-to-tail repeated sequence has six of the nine amino acids conserved.

A Ser-Xaa-Glu tripeptide sequence is abundant in the protein. About half (26 out of 55) of the serine residues and 80% (26 out of 33) of the glutamic acid residues occur in this constellation. It is possible that the Ser-Xaa-Glu sequence provides a signal for the N-acetylgalactosaminyltransferase, which is active in the first step in the synthesis of O-

Table 1. Amino acid composition of rat osteopontin. Comparison of experimentally determined values (7) with values predicted from the nucleic acid sequence of λ ROP11. The predicted composition is based on a signal-peptide cleavage site as indicated in Fig. 2

Residues	Determination Residues per 1000	Prediction	
		Residues per 1000	Residues per 295
Asn	_	24	7
Asp	_	142	42
Asx	147	166	49
Thr	44	41	12
Ser	153	186	55
Gln	·	51	15
Glu		112	33
Glx	184	163	48
Pro	56	48	14
Gly	32	20	6
Ala	75	68	20
Cys	_	0	0
Val	49	44	13
Met		7	2
Ile	23	24	7
Leu	61	54	16
Tyr	17	17	5
Phe	21	20	6
His	45	51	15
Lys	58	58	17
Arg	33	31	9
Trp	—	3	1

glycosidically linked oligosaccharides. The large number of these oligosaccharides in the protein is consistent with the abundance of Ser-Xaa-Glu sequences. There is only one Asn-Xaa-Ser sequence in the protein, representing the attachment site for N-glycosidically linked oligosaccharides (34).

The cDNA has a noncoding 3' sequence that contains four potential polyadenylylation signals (35). Polyadenylylation occurs at two sites, 14 or 17 nucleotides downstream of the last AATAAA sequence (indicated by arrowheads in Fig. 2). A similar variation in polyadenylylation sites has been reported for bovine prolactin mRNA (36) and mouse ribosomal protein L30 mRNA (37).

Cell Adhesion-Promoting Activity of Rat Osteopontin. The observation of the cell-binding sequence in osteopontin prompted us to study binding of the osteosarcoma cells (ROS 17/2.8) to purified osteopontin. The cells bound to and spread on surfaces coated with the protein. No attachment of cells was observed to plastic or to plastic surfaces coated with bovine serum albumin. The osteosarcoma cells similarly bound to and attained a spread morphology on surfaces coated with fibronectin.

To examine if the observed cell-attachment activity of osteopontin is contained in the Gly-Arg-Gly-Asp-Ser sequence, we used synthetic peptides in competitive inhibition experiments (Fig. 3). A Gly-Arg-Gly-Asp-Ser-Pro synthetic peptide, designed from the amino acid sequence of the cell-binding region in fibronectin, efficiently inhibited the binding of cells to osteopontin. A 50% inhibition of binding was obtained at 25 μ M peptide concentration. No inhibition of cell attachment was observed with a similar Gly-Arg-Gly-Glu-Ser-Pro peptide. This inactive peptide is different from the active peptide in a single position, where an aspartic acid has been replaced with a glutamic acid.

The predicted osteopontin sequence also contains the inverted cell-binding sequence Asp-Gly-Arg as part of the palindromic sequence Asp-Gly-Arg-Gly-Asp. It has been reported that the inverted sequence Ser-Asp-Gly-Arg is a



FIG. 3. Inhibition of ROS 17/2.8 cell adhesion to substratum coated with rat osteopontin. A suspension of cells (3×10^4) was added to coated microtiter wells in the presence of the indicated concentrations of synthetic peptides. •, Gly-Arg-Gly-Asp-Ser-Pro; \circ , Gly-Arg-Gly-Gly-Gly-Ser-Pro; \diamond , Lys-Pro-Ser-Asp-Gly-Arg-Gly; and \blacktriangle , Val-Pro-Asp-Gly-Arg-Gly-Asp-Ser-Leu-Ala-Lys. After incubation for 1 hr at 37°C, nonbound cells were rinsed away. The number of attached cells was determined and is shown as a percentage of cells added.

potent inhibitor of the attachment of baby hamster kidney (BHK) cells to fibronectin (38). However, the reversed cell-binding sequence does not appear to participate in the attachment of ROS 17/2.8 cells to osteopontin, since the synthetic peptide Lys-Pro-Ser-Asp-Gly-Arg-Gly was without effect in the competition assay.

A Val-Pro-Asp-Gly-Arg-Gly-Asp-Ser-Leu-Ala-Lys synthetic peptide, designed to mimic the osteopontin cell-binding sequence, also inhibited the adhesion of rat osteosarcoma cells to osteopontin. A 250 μ M solution of the undecamer was necessary to obtain 50% inhibition of cell attachment. Compared to the shorter active peptide Gly-Arg-Gly-Asp-Ser-Pro, a 10-fold higher concentration of the larger peptide was necessary for obtaining 50% inhibition. Possibly the undecamer assumes a conformation that mimics the cellbinding site in osteopontin less efficiently than does the shorter hexamer peptide. In support, it is known that the conformation of a specific region in a short peptide can be different from the conformation of the corresponding short region in the native protein (39). The results of these competition experiments strongly suggest that the Arg-Gly-Asp sequence is responsible for the cell adhesion-promoting activity of osteopontin.

Partial Amino Acid Sequence Analysis. Verification of the deduced amino acid sequence was obtained by sequence analysis of a fragment of osteopontin. Initial attempts to determine the sequence of intact osteopontin failed, possibly depending on a blocked N terminus or a result of partial cleavage of some of the molecules. Cyanogen bromide cleavage of osteopontin yielded three peptides, consistent with the presence of two methionine residues (residues 85 and 178 in Fig. 2) in the sequence deduced from the cDNA. The peptide with cell-binding activity was identified, and its amino acid sequence was determined. The N-terminal 14 amino acid residues had a sequence that agreed with that assigned from the cDNA analysis (underlined with a broken line in Fig. 2).

DISCUSSION

The present study shows that osteopontin belongs to a group of extracellular matrix proteins—e.g., fibronectin and vitronectin—that exert their cell-binding functions via an Arg-Gly-Asp sequence. An interesting goal should be identification of the cell-surface receptor that recognizes the

Biochemistry: Oldberg et al.

Arg-Gly-Asp sequence in osteopontin. The cell-surface receptor that recognizes fibronectin has been identified by affinity chromatography on immobilized fibronectin (40) and by immunological techniques (41). The fibronectin receptor is a 140-kDa membrane glycoprotein that specifically interacts with the Arg-Gly-Asp sequence in fibronectin but not with the same tripeptide sequence in vitronectin. Affinity chromatography of detergent-solubilized cells on immobilized vitronectin gives a vitronectin-specific cell-surface receptor (42). The vitronectin receptor is different from the fibronectin receptor and is composed of 115-kDa and 125-kDa polypeptide chains (42). Although fibronectin, vitronectin, and osteopontin interact with cells via the same Arg-Gly-Asp cell-binding sequence, it is possible that each cell-binding protein is recognized by a unique cell-surface receptor. The obvious function of a putative osteopontin cell-surface receptor is to provide bone cells with the ability to interact with their surrounding calcified matrix. The interaction between a unique cell-surface receptor and osteopontin could be important for guiding cells to areas of ossification and also could be involved in the development of the phenotype of bone cells.

In addition to its binding to a cell-surface receptor, fibronectin also interacts with collagen (43) and proteoglycans (44). It is not known whether osteopontin binds to these extracellular matrix components that are also present in the bone matrix. It is known, however, that osteopontin binds extremely tightly to hydroxyapatite (unpublished data), which is the most abundant component in bone calcified matrix. What region or regions in the osteopontin molecule are involved in this binding is not known. The bone-specific protein osteocalcin interacts with hydroxyapatite most likely via γ -carboxyglutamic acid residues (45). A potential region for mineral binding in osteopontin, containing numerous carboxyl groups, is the region consisting of nine consecutive aspartic acid residues. Perhaps this region can assume a conformation with similar carboxyl-group arrangements as the γ -carboxyglutamic acid residues in osteocalcin and, thus, bind to hydroxyapatite via a similar mechanism. Another possibility is that the numerous sialic acid-containing oligosaccharides confer affinity for hydroxyapatite to osteopontin.

We are much obliged and grateful to Drs. Gideon Roden and Robert Majeska for providing the ROS 17/2.8 osteosarcoma cell line, Dr. Tomas Brodin for the rat peritoneal fibroblasts, Drs. Michael Pierschbacher and Erkki Ruoslahti for providing the synthetic peptides, Dr. Åke Lundwall for help with the computer analysis, and Dr. Johan Stenflo for performing the amino acid sequence analysis. This work was supported by grants from the Swedish Medical Research Council (07478 and 05668), Folksams yrkesskadors stiftelse, Österlunds stiftelse, and the Medical Faculty, University of Lund.

- Linde, A., Jontell, M., Lundgren, T., Nilson, B. & Svanberg, V. (1983) J. Biol. Chem. 258, 1698-1705.
- Termine, J. D., Kleinman, H. K., Whitson, S. W., Conn. K. M., McGarvey, M. L. & Martin, G. R. (1981) Cell 26, 99-105.
- Pan, L. C. & Price, P. A. (1985) Proc. Natl. Acad. Sci. USA 82, 6109-6113.
- 4. Herring, G. M. & Kent, P. W. (1963) Biochem. J. 89, 405-414.
- Fisher, L. W., Whitson, S. W., Avioli, L. V. & Termine, J. D. (1983) J. Biol. Chem. 258, 12723-12727.
- 6. Franzén, A. & Heinegård, D. (1985) Biochem. J. 232, 715-724.
- 7. Franzén, A. & Heinegård, D. (1985) in Chemistry and Biology

of Mineralized Tissues, ed. Butler, W. T. (EBSCO, Birmingham, England), pp. 132-141.

- Höök, M., Rubin, K., Oldberg, Å., Öbrink, B. & Vaheri, A. (1977) Biochem. Biophys. Res. Commun. 79, 726-733.
- Rubin, K., Höök, M., Öbrink, B. and Timpl, R. (1981) Cell 24, 463–470.
- Terranova, V. P., Rohrbach, D. H. & Martin, G. R. (1980) Cell 22, 719-726.
- Hayman, E. G., Pierschbacher, M. D., Öhgren, Y. & Ruoslahti, E. (1983) Proc. Natl. Acad. Sci. USA 80, 4003– 4007.
- 12. Pierschbacher, M. D. & Ruoslahti, E. (1984) Nature (London) 309, 30-33.
- Pierschbacher, M. D., Ruoslahti, E., Sundelin, J., Lind, P. & Peterson, P. A. (1982) J. Biol. Chem. 257, 9593-9597.
- Suzuki, S., Oldberg, Å., Hayman, E. G., Pierschbacher, M. D. & Ruoslahti, E. (1985) *EMBO J.* 4, 2529-2524.
- Doolittle, R. F., Watt, K. W. K., Cottrell, B. A., Strong, D. D. & Riley, M. (1979) Nature (London) 280, 464–468.
- Gartner, K. T. & Bennett, J. S. (1985) J. Biol. Chem. 260, 11891-11894.
- Rodan, G. A. & Majeska, R.-J. in Limb Development and Regeneration, eds. Kelley, R. O., Goetinck, P. F. & Mac-Cabe, J. A. (Liss, New York), Part B, pp. 249–259.
- Brodin, T., Thurin, J., Strömberg, N., Karlsson, K. A. & Sjögren, H. O. Eur. J. Immunol., in press.
- Bourdon, M. A., Oldberg, Å., Pierschbacher, M. D. & Ruoslahti, E. (1985) Proc. Natl. Acad. Sci. USA 82, 1321-1325.
- 20. Gubler, U. & Hoffman, B. J. (1983) Gene 25, 263-269.
- Schwarzbauer, J. E., Tamkun, J. W., Lemischka, I. R. & Hynes, R. O. (1983) Cell 35, 421–431.
- 22. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- 23. Young, R. A. & Davies, R. W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194-1198.
- Oldberg, Å. & Ruoslahti, E. (1986) J. Biol. Chem. 261, 2113-2116.
- 25. Staden, R. (1982) Nucleic Acids Res. 10, 27-30.
- 26. Landegren, U. (1984) J. Immunol. Methods 67, 379-388.
- 27. Heinegård, D. (1977) J. Biol. Chem. 252, 1980-1989.
- 28. Dahlbäck, B., Lundwall, Å. & Stenflo, J. (1985) J. Biol. Chem., in press.
- 29. Von Heijne, G. (1983) Eur. J. Biochem. 133, 17-21.
- 30. Kozak, M. (1984) Nucleic Acids Res. 12, 857-870.
- 31. Blobel, G. & Dobberstein, B. (1975) J. Cell Biol. 67, 835-851.
- National Biomedical Research Foundation (1985) Protein Identification Resource (Natl. Biomed. Res. Found., Washington, DC 20007).
- 33. Staden, R. (1982) Nucleic Acids. Res. 10, 2951-2961.
- 34. Hughes, R. C. (1973) Prog. Biophys. Mol. Biol. 26, 189-268.
- 35. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) Cell 41, 349-359.
- Sasavage, N. L., Smith, M., Gillam, S., Woychick, R. P. & Rottman, R. M. (1982) Proc. Natl. Acad. Sci. USA 79, 223-227.
- 37. Wiedemann, L. M. & Perry, R. P. (1984) Mol. Cell. Biol. 4, 2528-2528.
- Akiyama, S. K. & Yamada, K. M. (1985) J. Biol. Chem. 260, 4492-4500.
- Kabsh, W. & Sanders, C. (1984) Proc. Natl. Acad. Sci. USA 81, 1075-1078.
- Pytela, R., Pierschbacher, M. D. & Ruoslahti, E. (1985) Cell 40, 191-198.
- 41. Brown, P. J. & Juliano, R. L. (1985) Science 228, 1448-1451.
- 42. Pytela, R., Pierschbacher, M. D. & Ruoslahti, E. (1985) Proc. Natl. Acad. Sci. USA 82, 5766-5770.
- 43. Engvall, E. & Ruoslahti, E. (1977) Int. J. Cancer 20, 1-5.
- 44. Oldberg, A. & Ruoslahti, E. (1982) J. Biol. Chem. 257, 4859-4863.
- 45. Poser, J. W. & Price, P. A. (1979) J. Biol. Chem. 254, 431-436.