Identification of the N-Linked Oligosaccharide Sites in Chick Corneal Lumican and Keratocan That Receive Keratan Sulfate*

(Received for publication, September 9, 1997, and in revised form, May 5, 1998)

Jane R. Dunlevy‡, Peter J. Neame, Jean-Paul Vergnes§, and John R. Hassell

From the Shriners Hospital for Children-Tampa, University of South Florida College of Medicine, Tampa, Florida 33612 and the §University of Pittsburgh School of Medicine, Department of Ophthalmology, Pittsburgh, Pennsylvania 15213

The corneal stroma contains an extensive matrix consisting of collagen and proteoglycans that interact with each other to produce a transparent structure. The proteoglycan constituents of the corneal stroma were originally designated as chondroitin/dermatan sulfate or keratan sulfate (KS)1 proteoglycans, but a number of these core proteins have since been cloned. They are now recognized as decorin (1), which contains a single chondroitin/dermatan sulfate side chain and lumican (2–4), osteointude factor (5, 6), or osteoglycin (7) and keratocan (5), which contain KS side chains. In adult chicken corneas, a portion of the population of decorin also receive KS to make a decorin isoform containing both chondroitin/dermatan sulfate and KS (8). KS is a repeating disaccharide of glucosamine and galactose that bears sulfate esters. In cornea, KS is N-linked to an Asn residue of an NX(S/T) site through a complex oligosaccharide linkage region. Only some sites receive KS, whereas others remain in a high mannose form. To determine whether the attachment of KS was biased toward specific sites, we isolated tryptic-digested KS-containing fragments of chick corneal proteoglycans and sequenced the peptides. Results showed that all of the peptides sequenced aligned to the deduced amino acid sequence of either chick lumican or chick keratocan at the first, third, and fourth potential N-linked sites. Sites 1 and 4 in lumican and keratocan are in a homologous location. By analogy with the structure of ribonuclease inhibitor (a Leu-rich repeat containing protein), the KS chains would extend outward on the outer face of a horseshoe-like structure. The amino acid sequences surrounding the potential N-linked sites were also compared. Sites receiving KS tend to have a higher occurrence of aromatic residues, in particular Phe, located within 3 amino acids of NX(S/T). These conserved Phe residues may have a role in the conversion of high mannose N-linked oligosaccharides to polygalactosamine and/or keratan sulfate.

Corneal proteoglycans have chondroitin/dermatan and keratan sulfate and KS (8). The nucleotide sequence(s) reported in this paper has been submitted with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Shriners Hospital for Children-Tampa, 12502 N. Pine Dr., Tampa, FL 33612.

§ To whom correspondence should be addressed: Shriners Hospital for Children-Tampa, 12502 N. Pine Dr., Tampa, FL 33612.

This paper is available on line at http://www.jbc.org

* This work was supported by National Institutes of Health Grants EY08104, EY06623, and EY08098. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

EXPERIMENTAL PROCEDURES

Isolation of cDNA Clone to Chick Keratocan—A previously prepared chick corneal cDNA library (2) was screened using a 32P-radioabeled cDNA clone to bovine keratocan (5) obtained by PCR. Briefly, the bovine keratocan clone was generated by reverse transcription-PCR using DNaase-treated RNA isolated from cultured bovine corneal fibroblasts and the following primers: reverse transcription primer 5’-GACTTTAGACTGCAAGATGGCG-3’; forward primer 5’-TTGGAAGATGAGCTCCCTTGAC-3’; reverse primer 5’-ATCCGAGGAGGTTACCGGAG-ATG-3’. The resulting 491-bp product was subcloned into the GenHunter vector (GenHunter Corp.) and sequenced to confirm its identity as bovine keratocan. The high specific activity PCR probe was generated in a standard 20-cycle PCR reaction containing 0.1 ng/μl plasmid DNA, forward and reverse keratocan primers, 20 μM dATP, dGTP, and dTTP, and 0.6 μM [α-32P]dCTP, 5000 Ci/nmol (NEN Life Science Products). Screening 140,000 plaque-forming units of the library under low stringency conditions yielded 250 positive clones; 74 phage plaques were extracted followed by PCR to determine which clones were full-length. Briefly, 5 μl of extracted phage was lysed in 35 μl of TE, 0.1% Tween 20 (GenHunter Corp.), boiled 5 min, and centrifuged 2
in min; the supernatant was collected, and 2.0 μl was used in a 20-μl, 30–40-cycle, PCR reaction with the T3 primer (Stratagene) and keratocan reverse primer. The five largest clones, determined by PCR, were used for the second library screen; positive clones were selected, and the pBluescript plasmid was rescued from the Uni-Zap XR vector (Stratagene) as a standard. The KS containing fractions were pooled, dialyzed against deionized water, and lyophilized to dryness.

**Identification of Keratan Sulfate Attachment Sites—KS Sites in Lumican and Keratocan**

A portion of the KS peptides were further digested with keratanase, as determined as above, were pooled, dialyzed against deionized water, and lyophilized to dryness. The proteoglycans were extracted at 4 °C for 16 h in 10 volumes (w/v) of 4 M guanidine HCl containing proteolytic inhibitors. The extract was dialyzed against 6 M buffered urea) and applied to a column (2.5 × 3 cm) of Superose 12 (Amersham Pharmacia Biotech) equilibrated and eluted in the same solvent. The tubes containing the KS peptides, determined by homology with chick decorin, were identified by the DMB assay, and the amount was calculated using KS (Seikagaku America) as a standard. The tubes containing proteoglycans were purified by application to a Vydac C-18 reversed phase column (30 × 2.1 mm) in 0.1% trifluoroacetic acid. Elution was with a linear gradient of acetonitrile (0–70% over 45 min) at a flow rate of 0.2 ml/min, and the eluate was monitored at 220 nm. Glycopeptide containing samples were compared with a buffer, and enzyme containing blank and peaks unique to the peptide containing sample were analyzed by Edman degradation on an Applied Biosystems/Perkin-Elmer 477A sequencer with on-line detection of phenylthiohydantoin-derivatives. Samples exhibited a prominent peak of Chaps which eluted later than the peptides. As is typical for glycopeptides, peaks were often broad and heterogeneous.

**RESULTS**

140,000 recombinants of a Uni-Zap XR cDNA chick corneal fibroblast library were screened using a high specific activity 32P-bovine keratocan probe generated by PCR. Full-length clones were determined using phage lysis and PCR with the T3 primer which anneals to the vector and the bovine keratocan reverse primer. PCR-selected phage clones were purified and rescued in the pBluescript plasmid from the Uni-Zap vector. Two putative clones, which were similar in insert size, were sequenced. The first clone consisted of a 2455-bp cDNA sequence containing a 138-bp 5'-untranslated region, a 353 amino acid open reading frame, and a 1257-bp 3'-untranslated region including the polyadenylated tail (Fig. 1). The deduced amino acid sequence of the full-length clone contains a signal peptide, underlined, estimated by the “-3, -1” rule (16), of 21 residues in length with a cleavage site between the Thr and Arg residues (Fig. 1). The sequence contains 7 Cys residues, one in the signal peptide, four near the N terminus and two near the C terminus. In addition, there are five potential N-linked oligosaccharide attachment sites, determined by the Asn-Xaa-(Ser/Thr) consensus sequence, located between the N- and C-terminal Cys residues. Although the polyadenylated tail was sequenced (data), the highly conserved AATAAA signal was not present 11–30 nucleotides upstream from the site of poly(A) addition. However, a similar sequence, AAATAA, was found located within this region. The second cDNA clone contained a 2456-nucleotide cDNA sequence that was identical to the sequence of the first clone except it contained additional 3'-untranslated sequence (not shown) and lacked the 5' 496 nucleotides which encoded for the N-terminal 119 amino acids. Interestingly, PRELP, another member of the Leu-rich proteoglycan (LRP) family, has been shown to have several different 3'-untranslated regions which accounts for the mRNA sizes of 1.7, 4.6, and 6.7 kilobase pairs (17).

A comparison of the deduced amino acid sequence of the first cDNA clone to sequences in GenBank revealed the highest homology to bovine keratocan, accession number U48360 (Fig. 2). There are only two regions where major variations exist, in the signal peptide and in the C-terminal domain at amino acids 306–320. The remainder of the protein is highly conserved with a 72.3% amino acid identity between chick and bovine keratocan. In comparison between chick and bovine species, keratocan was more conserved than lumican but not as conserved as decorin. Alignment of keratocan with other members of the LRP family revealed that both chick and bovine keratocan have the highest homology with human PRELP (18), a 53.3 and 54.5% amino acid identity, respectively. Chick keratocan had a slightly higher amino acid identity with bovine fibromodulin (19) than with chick lumican (2) and had a much lower homology with chick decorin (1), 39.8, 37.7, and 26.8%, respectively. It appears that keratocan, PRELP, fibromodulin, and lumican are all within the same gene family. This family can be divided further into two subgroups where keratocan and PRELP share >5% homology and lumican and fibromodulin share >45% homology between the two.

Chick keratocan contains 11 Leu-rich repeats, all of which are of the B-type (20), which is defined by a highly conserved Asn residue in the 10th position of the repeat. The Leu-rich repeat as well as the surrounding amino acids follow a pattern in which certain residues within each repeat are conserved at
the same position. Therefore, the first 9 Leu-rich repeats of chick keratocan can be organized into equally spaced motifs and units as shown in Fig. 3. The motifs 1, 2, and 3 are joined tandemly to form a unit. Each of the motifs has unique spacing and a unique Leu-rich repeat consensus sequence as follows: motif 1 has a spacing of 24 amino acid residues, motif 2 is the longest at 25–26 residues, and motif 3 is the shortest with 20–21 amino acids. This arrangement produces a triplet where the three motifs are aligned tandemly, and this arrangement is repeated three times. In addition to each motif having a unique consensus sequence within the Leu-rich repeat, there are also sequence identities that surround these repeats. For instance, Pro residues in position 16 of motif 1 and in positions 16 and 20 of motif 3 may give the motif a specific secondary structure. There are also five Phe residues, in motif 1 at position 20 and in motif 2 at position 22, that are not only highly conserved (see Fig. 2) but are located 5–6 amino acids N-terminal of a Leu-rich repeat. In addition to the 9 Leu-rich repeats shown in Fig. 3, there are two additional repeats that are present in chick keratocan. They are excluded from the units because their spacing is altered, and Leu-repeat 11 is located between the two C-terminal Cys residues making it structurally unique.

**Fig. 1.** cDNA and deduced amino acid sequence of chicken keratocan. The nucleotide sequence is listed with numbering on the left side of the figure, and the deduced amino acid sequence is listed below with numbering on the right side of the figure. The signal peptide sequence is underlined, Cys are marked (D), and the potential O-linked oligosaccharide sites are indicated (O).
The Asn residues in chick corneal proteoglycans that receive KS were also determined. Adult chicken corneas were extracted with 4 M guanidine in the presence of protease inhibitors, and the corneal KSPG proteoglycans were purified using both molecular sieve and ion exchange chromatography as well as digestion with chondroitinase ABC according to previous studies (8). This KSPG stock was then used to determine the KS attachment sites. The KSPGs were digested with trypsin, fractionated over a Source 15Q FPLC column, and eluted with NaCl, and the elution position of KS was determined using the DMB assay (Fig. 4). The DMB-positive fractions 15–30 were pooled and further purified on a Superose 12 FPLC column (Fig. 5). There were three KS containing peaks after Superose 12. The peak at fraction 10 had a very low DMB reactivity and may contain partially digested core protein due to its larger size. The peak in tubes 21/22 contained a high concentration of contaminating peptides and very small sized DMB-positive material; therefore, only the fractions in the largest main peak, fractions 12–17, were pooled and used for further analysis including amino acid sequencing.

The experimentally determined sequences from the KS peptides were then aligned with the deduced sequence from the cDNA clones (Fig. 6). The sequences containing potential N-linked glycosylation sites, starting at the estimated trypsin digestion site are marked with arrows, and any amino acid sequence, obtained by Edman degradation, of the KS containing peptides is listed below. There were a total of 11 peptide fragments that were sequenced, six of which aligned to chick lumican at three different N-linked glycosylation sites as follows: site 1, amino acids 77–93 with KS attachment on Asn-91; site 3, residues 154–167 with Asn-165 receiving KS chains; site 4, amino acids 245–259 with KS attachment on Asn-257. Similarly, five peptides that were sequenced aligned with the deduced amino acid sequence of chick keratocan at three different N-linked glycosylation sites as follows: site 1, amino acids 73–96 with KS attachment on Asn-94; site 3, residues 207–225 with Asn-216 receiving KS chains; site 4, amino acids 249–267 with Asn-261 receiving KS. Although the peptide fragments sequenced matched to sites 1, 3, and 4 in lumican.
and keratocan, it is still possible that sites 2 and 5 receive KS and that we did not sequence these fragments due to either low abundance (KS attachment at these sites is a rare event) or that the regions surrounding these sites were trypsin-resistant (in which the site would be contained in a large fragment such as the peak in fraction 10 after Superose 12 (Fig. 5)).

Although both proteoglycans receive KS attachment on sites 1, 3, and 4, only sites 1 and 4 are located in a homologous region of the core protein (see Fig. 8). Several peptides contained one or two residues that were undetermined, indicated by the X, during cycle sequencing which may be due to carbohydrate linked to the core protein at this site. In addition, there was one discrepancy between the deduced amino acid sequence and the peptide sequencing. The two peptides sequenced which aligned to lumican, residues 245–259, contain a Ser at position 247, whereas the cDNA deduced amino acid sequence indicates that this position should contain a Thr residue. The data on the amino acid sequencing were clear, and it is possible that there was a single nucleotide change during library construction or that this difference is due to a polymorphism. All other peptide fragments matched lumican and keratocan exactly. None of the peptides aligned to other published sequences, and no peptide sequence was obtained that did not match either lumican or keratocan.

The sequence flanking the NX(S/T) sites containing KS in lumican and keratocan were compared with sequence flanking the NX(S/T) sites that did not contain detectable KS (Fig. 7A). One striking difference in the sequences was the presence of aromatic residues surrounding most of the NX(S/T) sites that contain KS. The sequences surrounding the KS attachment sites contained 7 aromatic residues compared with only 2 aromatic residues surrounding the sites that were not found to contain KS. In particular, most of the sites that contain KS have a Phe residue present at a position within 3 amino acids N-terminal to the Asn residue. The only exception was site 3 in lumican which did not contain any aromatic residues in this region.

**DISCUSSION**

The results of this study show that chick keratocan has a message size of 2.45 kilobase pairs that encodes for a 353-amino acid protein which has high homology with bovine keratocan (72.3%). As is typical of proteins in the LRP family, chick keratocan contains 11 Leu-rich repeats that can be arranged into three units comprised of three motifs. This arrangement reveals the unique consensus sequence for each units’ Leu-rich repeat and the amino acids surrounding it in the motif.

This study also shows that both chick lumican and chick keratocan have three out of five potential N-linked oligosaccharide sites substituted with KS chains. N-Linked oligosaccharides at sites 1, 3, and 4 in both chick lumican and keratocan are modified to KS. There were, however, some KS containing peptides that were exceptionally large (Fig. 5) that we did not attempt to characterize, and they may contain additional sequences. Previous studies on KSPGs from bovine cornea indicated that keratocan has three KS chains, whereas lumican and osteoglycin contained only one KS chain (12). In the present study, all the sequences we detected aligned with either lumican or keratocan. It may be that chick corneal osteoglycin is present in too low abundance to be detected in the material that was characterized. In support of this, bovine KS core proteins were found to be present at a 3:6:2 ratio, for keratocan, lumican, and osteoglycin, respectively (4, 12).

Comparison of N-linked sites that have KS chains to sites that were not found in the KS-containing peptide fractions show that KS-containing sites tend to have a nearby aromatic residue, usually a Phe. More specifically, sites 1, 3, and 4 of keratocan and sites 1 and 4 of lumican all have Phe located within 3 residues N-terminal to the Asn. Conversely, all of the sites that were not found to contain KS chains lack Phe in this region. In support of the possible role of aromatic residues potentially enhancing KS attachment, fibromodulin, which can have KS chains on the first 4 out of 5 N-linked oligosaccharide sites (21), has at least one aromatic residue located near (within 5 amino acids) the Asn in the NX(S/T) consensus site (Fig. 7B). Furthermore, sites 1 and 4 in fibromodulin contain Phe within 3 amino acids of the Asn with KS attached. There are exceptions to this trend, such as site 3 in chick lumican lacks an aromatic residue near the Asn, yet supports KS attachment.

**FIG. 5.** KS containing fractions after Source 15Q FPLC were pooled and further purified on a Superose 12 FPLC column. The protein content for each tube is indicated by the O.D. at 280 nm on the left-handed y axis. A DMB assay was used to determine which fractions contained KS as indicated on the right-handed y axis. The bar on the top of the graph shows which tubes were pooled and used for amino acid sequencing.

**FIG. 6.** Potential N-linked glycosylation sites for chick keratocan and lumican. The top line marked with an arrow represents the deduced amino acid sequence, and the lines below it represent residues determined by amino acid sequencing KS containing fragments of corneal KSPGs. The N-linked glycosylation consensus sequence NX(S/T) is underlined.
whereas chick lumican site 2 contains a Tyr residue two amino acids N-terminal to the Asn but was not found in the GAG attachment fractions. Thus, the presence and proximity of aromatic residues, in particular Phe, near the Asn may enhance or facilitate KS attachment but is not required or necessary for this processing to occur.

Most of the Phe residues in close proximity to KS-substituted sites are also located 5–6 amino acids N-terminal to a Leu-rich repeat. This occurs with sites 1 and 4 in lumican, keratocan, and fibromodulin and may indicate that the potential role of these Phe residues is to support the secondary structure necessary for the β-sheet (i.e. at the top of the coil, see Fig. 8) as well as KS attachment. Regardless of the signals present in the core protein which may enhance or signal KS attachment, the developmental or tissue-specific expression of the transferases and sulfotransferases required for KS synthesis will regulate whether the core protein can be processed to the proteoglycan form.

The location of the KS attachment sites in lumican and keratocan were compared in a model based on the x-ray crystallography structure of porcine ribonuclease inhibitor, a protein comprised entirely of Leu-rich repeats (20, 22). A major feature of this model is that the Leu-rich repeats are coiled in a spiral with the Leu-rich regions stacked in a parallel β-sheet array. Fig. 8 shows all the potential N-linked sites and those substituted with KS in a two-dimensional model for lumican and keratocan. The first 10 Leu-rich repeats are contained in the central coiled domain between the N- and C-terminal globular domains. The 11th Leu-rich repeat for both lumican and keratocan is located between the two Cys within the C-terminal globular domain. The putative β-sheet, comprised of residues $L_XL_XL_XN$ within the Leu-rich repeat, is shown as the top third of the coil in which $X_L_X$ (i.e. amino acids YLY in repeat 1) are located at the top, exposed surface of the spiraled coil. Based on the structure of ribonuclease inhibitor, these coils would bend into a horseshoe-like structure (as implied in Fig. 8) with the Leu-rich repeats in closer proximity to each other thereby forming the inner, concave surface of the horseshoe. This hypothetical configuration also allows all three of the KS attachment sites in lumican and keratocan to extend outward on the outer, convex surface of the horseshoe structure. It is possible that the inner, Leu-rich, surface of the horseshoe-like structure is involved in protein-protein interaction, possibly collagen binding, whereas the outer surface, with the protruding GAG side chains, is involved in “space filling” between the collagen molecules.

Comparison of the GAG attachment sites within keratocan, lumican, and fibromodulin indicate that sites 1 and 4 in each of these proteoglycans are located in the same homologous region of the core protein.
of the core protein, and all of these proteoglycans support KS attachment at these sites. These two sites are located three residues N-terminal to a Leu-rich repeat which, according to our structural model shown in Fig. 8, allows the GAG chain to protrude outwards from the core protein. Site 3 which also receives KS chains is located in a different position to the Leu-rich repeat than sites 1 and 4. Lumican and fibromodulin have a homologous site 3 which is located at position 1 of a Leu-rich repeat. Keratocan has a unique site 3 which is located at position 9 within the Leu-rich repeat. Interestingly, this repeat is “unconventional” due to a Thr residue being present at the first position of the consensus LXXLXXN repeat instead of a Leu or Ile residue. This may represent a structure unique to keratocan. Despite these differences with site 3 between lumican and keratocan, our model in Fig. 8 indicates that the predicted three-dimensional structure would place site 3 in a similar location, and the GAG chains may also extend in the same direction so that lumican and keratocan may have a very similar three-dimensional appearance.

Acknowledgments—We are grateful to Pam Cornuet for outstanding technical assistance and Dr. Anna Plaas at the Shriners Hospital for Children-Tampa Unit for comments on the manuscript.

REFERENCES
1. Li, W., Vergnes, J. P., Cornuet, P. K., and Hassell, J. R. (1992) Arch. Biochem. Biophys. 296, 199–197
2. Blochberger, T. C., Vergnes, J. P., Hemper, J., and Hassell, J. R. (1992) J. Biol. Chem. 267, 347–352
3. Chakravarti, S., Stallings, R. L., SundarRaj, N., Cornuet, P. K., and Hassell, J. R. (1995) Genomics 27, 481–488
4. Funderburgh, J. L., Funderburgh, M. L., Brown, S. J., Vergnes, J. P., Hassell, J. R., Mann, M. M., and Conrad, G. W. (1993) J. Biol. Chem. 268, 11874–11880
5. Corpuz, L. M., Funderburgh, J. L., Funderburgh, M. L., Bottomley, G. S., Prokash, S., and Conrad, G. W. (1996) J. Biol. Chem. 271, 9759–9763
6. Madisen, L., Neubauer, M., Flowman, G., Rosen, D., Segarini, P., Dasch, J., Thompson, A., Ziman, J., Bentz, H., and Purich, A. F. (1990) DNA Cell Biol. 9, 303–309
7. Ujita, M., Shinomura, T., and Kimata, K. (1995) Gene (Amst.) 158, 237–240
8. Blochberger, T. C., Cornuet, P. K., and Hassell, J. R. (1992) J. Biol. Chem. 267, 20613–20619
9. Nilsson, B., Nakazawa, K., Hassell, J. R., Newsome, D. A., and Hascall, V. C. (1983) J. Biol. Chem. 258, 6056–6063
10. Nakazawa, K., Newsome, D. A., Nilsson, B., Hascall, V. C., and Hassell, J. R. (1983) J. Biol. Chem. 258, 6051–6055
11. Midura, R. J., and Hascall, V. C. (1989) J. Biol. Chem. 264, 1423–1430
12. Funderburgh, J. L., Funderburgh, M. L., Mann, M. M., and Conrad, G. W. (1991) J. Biol. Chem. 266, 14226–14231
13. Nakazawa, K., Hassell, J. R., Hascall, V. C., Lehmander, S., Newsome, D. A., and Krakemer, J. (1984) J. Biol. Chem. 259, 13751–13757
14. Farndale, R. W., Sayers, C. A., and Barrett, A. J. (1982) Connect. Tissue Res. 9, 247–248
15. Oike, Y., Kimata, K., Shinomura, T., Nakazawa, K., and Suzuki, S. (1980) Biochem. J. 191, 193–202
16. von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4680
17. Grover, J., Chen, X.-N., Kornberg, J. R., Beckley, A. D., and Roughley, P. J. (1996) Genomics 38, 109–117
18. Bengtsson, E., Neame, P. J., Heineg˚rd, D., and Sommarin, Y. (1995) J. Biol. Chem. 270, 25639–25644
19. Oldberg, A., Antonsson, P., Lindblom, K., and Heineg˚rd, D. (1989) EMBO J. 8, 2601–2604
20. Kobe, B., and Deisenhofer, J. (1993) Nature 366, 751–756
21. Plaas, A. H. K., Neame, P. J., Novens, C. M., and Reiss, L. (1990) J. Biol. Chem. 265, 20634–20640
22. Kobe, B., and Deisenhofer, J. (1995) Nature 374, 183–186
Identification of the N-Linked Oligosaccharide Sites in Chick Corneal Lumican and Keratocan That Receive Keratan Sulfate
Jane R. Dunlevy, Peter J. Neame, Jean-Paul Vergnes and John R. Hassell

J. Biol. Chem. 1998, 273:9615-9621.
doi: 10.1074/jbc.273.16.9615

Access the most updated version of this article at http://www.jbc.org/content/273/16/9615

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 22 references, 12 of which can be accessed free at http://www.jbc.org/content/273/16/9615.full.html#ref-list-1