PROTEINS: Structure, Function, and Bioinformatics 61:227–238 (2005)

# Structures of Vertebrate Hyaluronidases and Their Unique Enzymatic Mechanism of Hydrolysis

**Mark J. Jedrzejas[1]* and Robert Stern[2]**
[1]*Children's Hospital Oakland Research Institute, Oakland, California*
[2]*Department of Pathology, Comprehensive Cancer Center, School of Medicine, University of California, San Francisco, California*

***ABSTRACT*** **Human hyaluronidases (Hyals) are a group of five endo-β-acetyl-hexosaminidase enzymes, Hyal-1, -2, -3, -4, and PH-20, which degrade hyaluronan using a hydrolytic mechanism of action. Catalysis by these Hyals has been shown to follow a double-displacement scheme. This involves a single Glu residue within the enzyme, the only catalytic residue, as the proton donor (acid). Also involved is a carbonyl group of the hyaluronan (HA) *N*-acetyl-*D*-glucosamine as a unique type of nucleophile. Thus the substrate participates in the mechanism of action of its own catalysis. An oxocarbonium ion transition state is postulated, but there is no formation of a covalent enzyme–glycan intermediate, as found in most such reactions. The major domain is catalytic and has a distorted $(\beta/\alpha)_8$ triose phosphate isomerase (TIM) barrel fold. The C-terminal domain is separated by a peptide linker. Each Hyal has a different C-terminal sequence and structure, the function of which is unknown. These unique C-termini may participate in the additional function(s) associated with these multifunctional enzymes. Proteins 2005;61:227–238.**


**Key words: chondroitin; hyaluronan; hyaluronidase; mechanism of action; modeling; structure/function**

## INTRODUCTION

Hyaluronan (HA) is a simple unadorned glycosaminoglycan (GAG) linear polymer, built of a large number of repeating units of [*D*-glucuronic acid-β1,3-*N*-acetyl-*D*-glucosamine-β1,4-]$_n$. The enzymes that degrade HA are termed the hyaluronidases (Hyals) (Fig. 1). Degradation of HA is performed by the Hyals utilizing one of two mechanisms, either hydrolase or lyase, each of which involves cleavage of the β1,4-glycosidic bond. Mammalian Hyals (EC 3.2.1.35), members of glycosidase family 56,[1,2] are endo-β-acetyl-hexosaminidases that utilize hydrolysis in their catalysis of HA. On the other hand, the bacterial Hyals, which are also endo-β-acetyl-hexosaminidases, utilize a lyase mechanism, and are referred to as hyaluronate/hyaluronan lyases. Both groups of enzymes, hydrolases and lyases, also have, to a limited degree, the ability to degrade unsulfated chondroitin (Ch) and sulfated chondroitins (ChS). However this occurs at a much slower rate and is restricted to only certain sulfation patterns (Fig. 1). Chondroitins are polymeric GAGs differing from HA only in the anomeric configuration at the C4 atom of the *N*-acetyl-*D*-glucosamine, making it *N*-acetyl-*D*-galactosamine.

In the human genome, there are six known genes coding for Hyals: *hyal-1*, *hyal-2*, *hyal-3*, *hyal-4*, and *PH-20/Spam1* as well as a pseudogene *Phyal1* that is transcribed but not translated. The first three Hyal genes, *hyal-1–3*, are tightly clustered together on chromosome 3p21.3, and the latter three, *hyal-4*, *PH-20/Spam1*, and *Phyal1*, are clustered similarly on chromosome 7q31.3.[3] Hyal-1 and Hyal-2 are the two major human Hyals in somatic tissues. Hyal-2 degrades high molecular weight HA to an approximately 20 kDa product (~50 HA disaccharide units), whereas Hyal-1 can degrade high molecular weight HA to smaller products, predominantly tetrasaccharides. Hyal-3 protein product is expressed, but the precise nature of its activity is still being investigated (B. Triggs-Raine, personal communication). A product of the *hyal-4* gene, Hyal-4, also has chondroitinase activity (unpublished results). Human PH-20 or SPAM1 (SPerm Adhesion Molecule 1) (HPH-20) is necessary for fertilization, is commonly associated with testes, and also facilitates penetration of sperm through the cumulus mass in its passage to the ovum. A product of a seventh human gene has been proposed to have Hyal activity,[4] but its identity as a Hyal has not been explored and therefore will not be discussed further.

By degrading HA, the Hyals contribute to tight regulation of steady-state levels of HA deposition. But they are also crucial for preserving the biophysical properties of size-specific HA oligomers, an information-rich system

Abbreviations: 3D, three-dimensional; aa, amino acid; Ch, unsulfated chondroitin; ChS, chondroitin sulfate; BVHyal, bee venom Hyal; BPH-20, bovine PH-20; GAG, glycosaminoglycan; HA, hyaluronan, hyaluronic acid; HPH-20, human PH-20; Hyal, hyaluronidase; NMR, nuclear magnetic resonance; pdb, protein data bank (3D structure depository); TIM, triose phosphate isomerase.

Grant sponsor: National Institutes of Health (to MJJ); Grant number: AI44078; Grant sponsor: California Spinal Cord Injury Research Fund (to RS).

*Correspondence to: Mark Jedrzejas, Children's Hospital Oakland Research Institue, 5700 Martin Luther King, Jr. Way, Oakland, California 94609, USA. E-mail: mjedrzejas@chori.org

Received 29 October 2004; Revised 2 March 2005; Accepted 21 March 2005

Published online 15 August 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20592

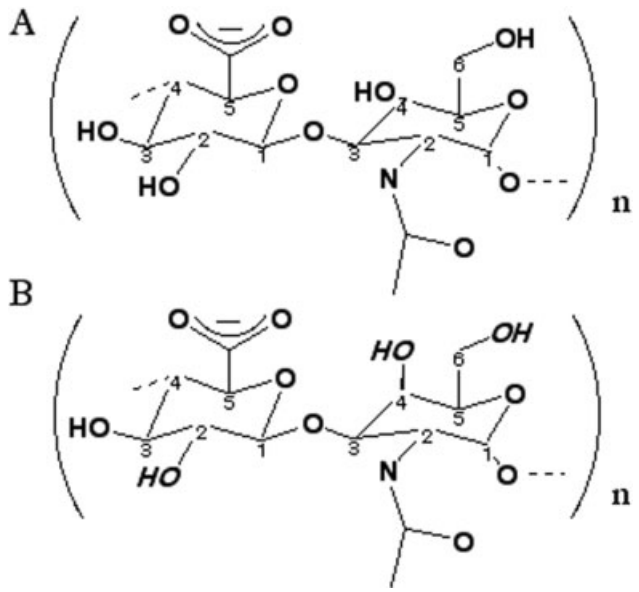


Fig. 1. Chemical structures of (**A**) hyaluronan and (**B**) unsulfated chondroitin and chondroitin sulfates. The structures differ only in the anomericity at the C4 position of the *N*-acetyl-*D*-glycosamine, *N*-acetyl-*D*-glucosamine in hyaluronan and *N*-acetyl-*D*-galactosamine in chondroitin. Potential sulfation sites for chondroitin structures are marked by italicized hydroxyl groups. Both glycans are substrates for the human hyaluronidase enzymes. For Hyal-1 and -2, HA is the primary substrate, but degradation of Ch/ChS also occurs, albeit at a slower rate, as documented *in vitro*. Hyal-4 is a chondroitinase as well as a degrader of HA.

associated with a wide range of biological activities. HA turnover by the Hyals is also critical in a number of important regulatory processes, ranging from embryonic development to wound healing (for review see ref. 3). The detailed understanding of the enzymatic mechanism(s) involved in HA degradation will certainly facilitate better understanding of the entire process of HA catabolism. The precise enzymatic mechanism of the vertebrate Hyals is still under investigation, but based on structural investigations of bee venom hyaluronidase (BVHyal) by Markovic-Housley and coworkers it is proposed to involve a double-displacement mechanism at the C1 anomeric carbon atom of the *N*-acetyl-*D*-glucosamine.[5,6] Such a mechanism results in net retention of the anomeric configuration of the substrate at the carbon C1 position.[7,8] There is no Hyal-covalent intermediate involved in such a process, as is the case for most of the other glycan hydrolases.[9]

## MATERIALS AND METHODS

### Sequences and Their Analyses

Sequence data for human hyaluronidases were retrieved from GenBank or other websites interfaced with the National Center of Biotechnology Information (NCBI). Protein or domain alignments were constructed using Clustal W 1.82.[10] Secondary structure elements were predicted using Protein Structure Prediction server (PSIPRED).[11] Fold recognition analyses were carried out using the Structure Prediction META server with the implemented 3Djury list (bioinfo.pl/Meta).[12,13]

### Three-Dimensional (3D) Modeling

Comparative/homology and *ab initio* modeling of the Hyals and their domains were carried out at the Robetta Structure Prediction Server (robetta.bakerlab.org).[14,15] The server returned ten models, the first nine representing the most populated clusters, and the tenth the lowest-energy structure outside those clusters. The models were assessed using the protein structure validation tools Protein Structure Analysis II (PROSA II),[16] VERIFY_3D,[17] and Protein Quality Prediction (ProQ).[18] The 3D structures and their properties were visualized with O.[19] The cartoon figures were obtained using PyMol software (www.pymol.org),[20] Ribbons,[21] and Grasp[22] programs.

The model of the Hyal-1 complex with an HA tetrasaccharide was obtained based on positioning the HA substrate in an identical Hyal-1 active-site cleft, as in the BVHyal and BHPH-20 structures in complex with HA tetrasaccharide [Protein Data Bank (PDB) (3D structure depository) code 1FCV].[6,23] All structures were aligned in 3D using the Distance Matrix Alignment (DALI) server (www.ebi.ac.uk/dali).[24] The molecular manipulations were performed with moleman package as implemented in O package.[19]

## RESULTS AND DISCUSSION

### Sequence Analysis

The protein sequences of all full-length human Hyal genes, *hyal-1–4*, and *PH-20*, either identified experimentally or taken from the human genome (available at www.ncbi.nlm.nih.gov) were used in the sequence-homology and structural studies. These were used to further elucidate the primary, secondary, and 3D structural properties of these proteins. It was presumed that such an approach would yield functional properties and mechanisms of action of these Hyals in greater detail. The sequences of the human Hyals described above (Hyal-1–4 and HPH-20) were found to be relatively uniform in their amino acid (aa) lengths, ranging from the shortest, Hyal-1 with 435 aa, to the longest, HPH-20 with 510 aa residues.

Fig. 2. Sequence of the human hyaluronidases. (**A**) Sequence alignment of Hyal-1–4, PH-20/SPAM1 and BVHyal with the known X-ray structure. The sequence alignment was performed using the Clastal W 1.82 software package.[25] The identity between all sequences varied from 33.1% between Hyal-3 and -4 to 41.2% between Hyal-4 and HPH-20. The identity with the sequence of the BVHyal enzyme (with known X-ray structure) in the aligned region ranges from 22.9% for PH-20 to 25.2% for Hyal-1. The catalytic Glu H-donor residue and the residues positioning the nucleophile/base of the substrate are strictly conserved (Table II) with the exception of the Cys263 residue of Hyal-4, which might reflect this enzyme's specificity for chondroitin. The types of aa residues are color-coded as follows: red, AVFPMILW small residue; green, STYHCNGQ hydroxyl, amine, or basic; blue, DE acidic; magenta, RHK basic; and gray, others. The conserved residues are marked as follows: (*) identical in entire column, conserved according to color scheme above, and exhibiting some semi-conserved substitutions. The proposed catalytic Glu H-donating residue is also marked with $\sigma$, whereas residues positioning the carbonyl nucleophile/base are marked with $\lambda$. The sequence of the portion of the BVHyal enzyme that was crystallized is underlined. (**B**) Schematic diagram of domain composition of human Hyals. Human Hyals are composed of two domains, a major catalytic domain followed by a C-terminal one of unknown function. These domains are connected by a presumably flexible peptide linker. The short segment at the extreme N-terminus is independent of the catalytic domain and assumes an $\alpha$-helical conformation.

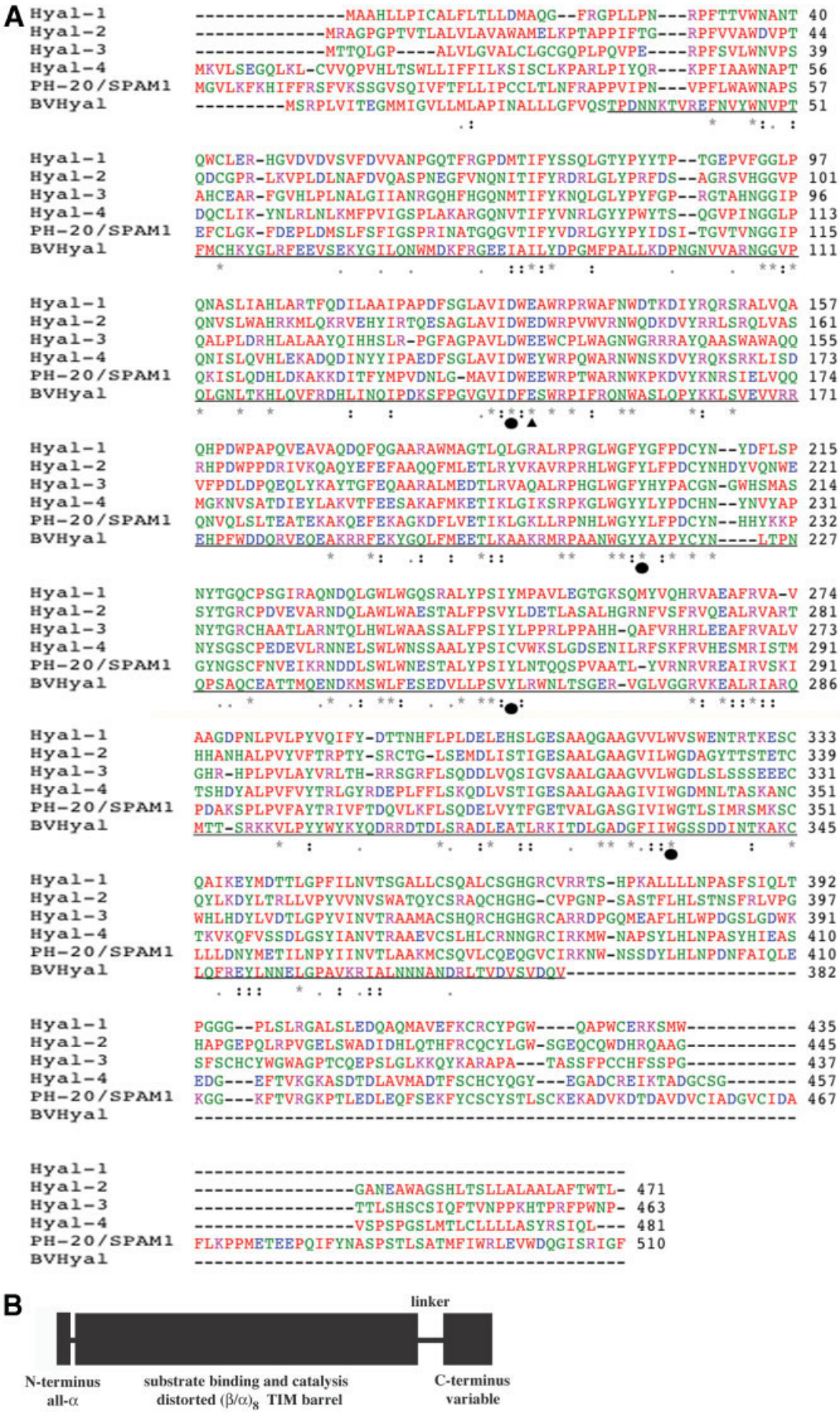


Figure 2.

The pairwise and multiple-sequence alignments of known human Hyals (Hyal-1–4 and HPH-20) performed using multiple sequence alignment Clastal W (1.82)[25] demonstrated that their sequence identities ranged from 33.1% (between Hyal-4 and Hyal-3) to 41.2% (between Hyal-4 and HPH-20). This degree of sequence conservation is indicative of the Hyals presumed common structural properties (Fig. 2). In addition, there are a number of absolutely conserved regions of aa sequence, suggesting the functional and mechanistic importance of these residues. The major differences appear to be located to the extreme N- and C-termini of all of these proteins.

The primary sequence alignments described above (Fig. 2) indicate that all Hyals, in addition to being homologous to one another, are also homologous to the bee venom Hyal (BVHyal),[6] bovine PH-20 (BPH-20) enzyme,[23] as well as other vertebrate Hyals (data not shown) (Fig. 2). The significance of the homology of human Hyals to BVHyal and BPH-20 enzymes, is that these are the only two Hyals that have known 3D structures, obtained by experimental X-ray crystal structure and by modeling, respectively.[6,23]

### Fold Recognition Studies of the Human Hyals

The fold recognition analyses were performed utilizing primarily the META server (bioinfo.pl/Meta).[12] All methods employed at the META server indicated consistently that all human Hyals are similar to one another at the primary, secondary, and tertiary levels. All Hyals were shown to be similar in fold to the BVHyal enzyme, for which a 3D X-ray structure has been elucidated (PDB code: 1FCQ). This similarity was also recognized by the BLAST and PDB-BLAST methods, with scores ranging from 2 $e^{-30}$ to 1 $e^{-37}$ for BLAST searches, and 1 $e^{-115}$ to 1 $e^{-118}$ for PDB-BLAST.[26] By the 3D-Jigsaw/Pcons3 consensus fold recognition predictor, highly significant scores of above 5 for all Hyals were obtained, well above the level of the highest-scoring false positives.[27] The BVHyal structure was consistently the top-scoring structure for all individual consensus fold recognition methods utilized in 3D-Jury, for which highly significant levels were reached.[13,28]

These results also supported the observation that all human Hyal enzymes are in the CAZY glycoside hydrolase family 56,[1,2] and classified by the SCOP database[29,30] as members of the (trans)glycosidase superfamily of enzymes C.1.8. Other members of such a SCOP superfamily include a catalytic domain of amylase[31] and type II chitinase enzymes.[32] Amylase is an endoglycosidase hydrolase that cleaves the α-bonds of another simple sugar polymer, glycogen. Chitinase is also an endoglycosidase, an endohexosaminidase cleaving β1,4-N-acetyl glycosaminidase bonds. Additional support for the structural and functional correspondence of all of these proteins comes from the excellent match between all predicted Hyal secondary structures and the actual secondary structure of the BVHyal obtained by X-ray crystallography (data not shown). Therefore, the fold recognition studies suggest that all human Hyals have 3D structures that are similar to one another, and also similar to that of BVHyal, and therefore presumably to all other vertebrate Hyals (data not shown).

### Homology and *Ab Initio* Modeling of 3D Structures for Hyal-1–4 and HPH-20

Although the human Hyal sequences share only moderate (22–25%) sequence identity with BVHyal, the fold recognition studies, as described above, demonstrate that the 3D structure of the BVHyal enzyme is the best and the only structural template available for modeling these proteins. The relatively low sequence identity is not reflective of a high degree of structural similarity among these Hyals. However, protein structures with such low levels of identity are often found appropriate and suitable for homology modeling.[33,34] In this regard, the fold recognition methods indicated unambiguously significantly higher structural conservation than that suggested using sequence identities. Due to the similarity to the BVHyal structure, there was a clear opportunity to model the structures of human Hyals primarily by homology. To accomplish such modeling, we chose to utilize the methodology implemented by the Robetta structure prediction server.[14] All models, including Robetta homology/comparative models, must be viewed with caution and compared to experimental structures obtained using either X-ray crystallography or nuclear magnetic resonance (NMR). However, the high structural homology with the BVHyal structure allowed the Robetta server to perform a reliable and consistent 3D structure model determination for Hyal-1–4 and HPH-20 enzymes. These enzymes are thought to be structurally representative of all vertebrate Hyals. By obtaining such models, we hoped to explore the mechanisms describing how the Hyals degrade HA, and possibly explain the known differences among the various human Hyals, such as why Hyal-1 degrades HA to tetrasaccharides whereas Hyal-2 degrades the same substrate to larger fragments of approximately 50 disaccharides or 20 kDa. For HPH-20 we had hoped to elaborate on the enzyme's ability to facilitate sperms' entry through the cumulus mass or zona pellucida into the human ovum. The insight into more functional and mechanistic details of Hyal-3 and Hyal-4 was also desired, as these details are lacking compared to other human Hyals.

All models obtained for Hyal-1–4 as well as for HPH-20 are of high quality and essentially identical to one another in the structure of their main portions. The structural models produced by the Robetta server were evaluated using objective protein-quality measures as implemented in PROSA II,[16] VERIFY_3D,[17] and ProQ.[18] The bases for these quality analyses were largely independent of those contained within the Robetta structure prediction, and therefore suited to identify the best, correct model out of the 10 models provided. As expected, one of the models produced by Robetta for each target, number 4, 1, 5, 5, and 3 for Hyal-1, -2, -3, -4 and HPH-20, respectively, performed best by all three methods, ProQ, PROSA II, and VERIFY_3D. These models (Figs. 3 and 4) provided predicted structures similar to that of BVHyal in their main,

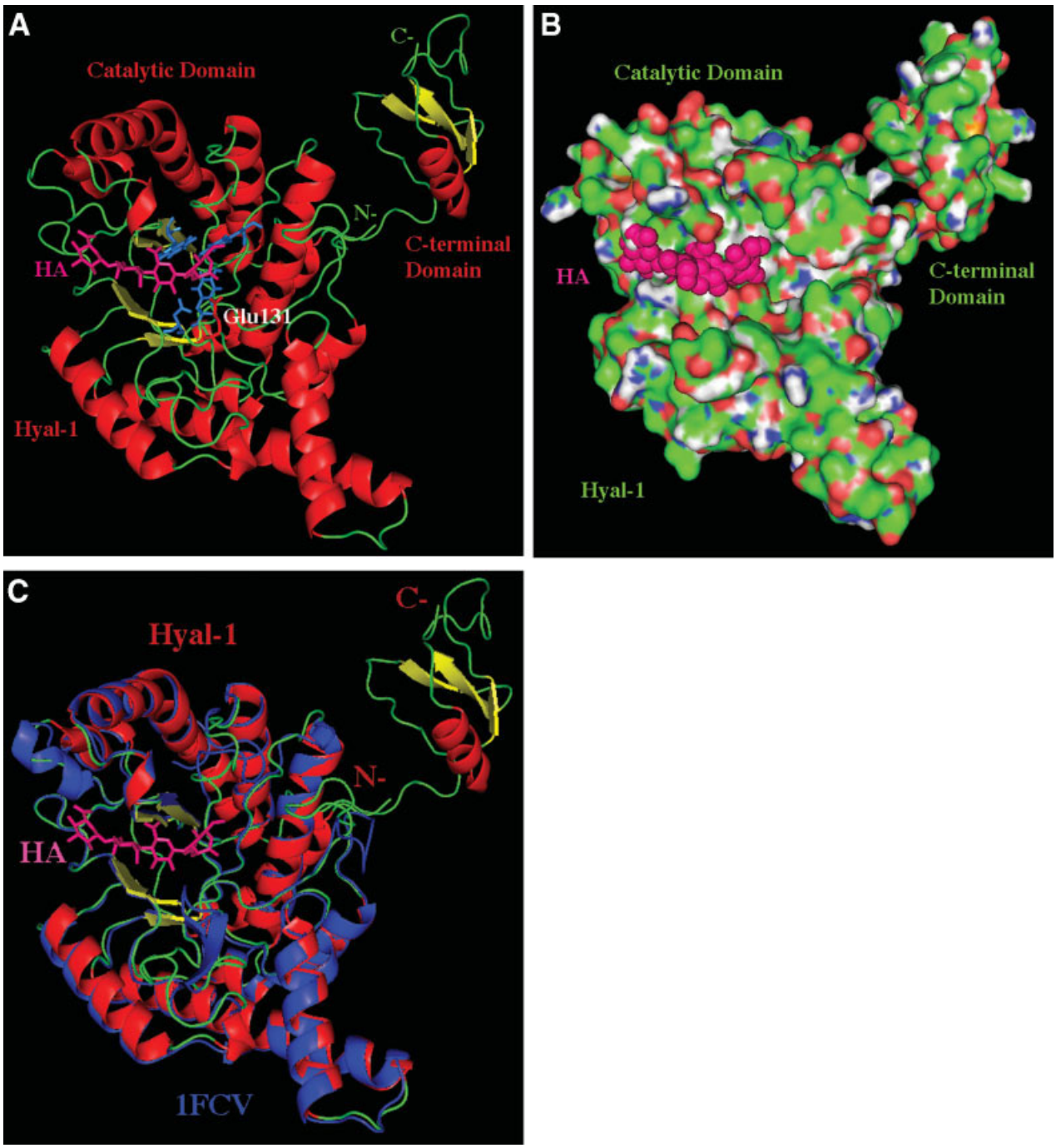


Fig. 3. Ribbon diagrams of 3D structural model of human Hyal-1. All three models were obtained using the Robetta server[14] and created based primarily on the comparative/homology modeling for the main, catalytic, domain utilizing the structure of bee venom hyaluronidase (PDB code 1FCQ).[6] The C-terminal domain was modeled using *ab initio* methods, and its primary, secondary, and tertiary structures are not similar to any molecule's 3D domain structure currently available, which implies a novel fold and possibly a novel function. There is a characteristic $(\beta/\alpha)_8$ TIM barrel fold of the main domain, which is the catalytic domain that supports substrate-binding and its hydrolytic degradation by a double-displacement retaining mechanism (described in the text). (**A**) Three-dimensional model of Hyal-1 degrading HA, primarily to tetrasaccharides. The molecule is color-coded by the secondary structure elements (helices in red, β-sheets in yellow, others in green). The bound HA molecule is located in the HA-binding cleft and is depicted in ball-and-stick fashion and colored in purple. The catalytic residue, Glu131, is also shown in ball-and-stick fashion colored in red (labeled); other residues that position the carbonyl of the acetylamido group of HA are also shown and are colored in blue. The catalytic and C-terminal domains, as well as the N- and C-termini, are shown and labeled. (**B**) Comparison of 3D structure of Hyal-1 and BVHyal enzymes. The Hyal-1 molecule is shown in the same color and orientation as in **A** (including HA). The BVHyal enzyme structure is overlaid on the structure of the Hyal-1 model and colored in blue. The positions of the catalytic Glu and carbonyl positioning residues are essentially identical for both structures (data not shown). The BVHyal does not have the C-terminal domain. (**C**) Surface of the Hyal-1 molecule and a tetrasaccharide HA substrate bound in the cleft. The orientation of the molecule is similar to that in **A-B**. The protein surface is colored by the atomic element: C, green; N, blue; O, red; S, yellow. The surface of the HA tetrasaccharide bound to the enzyme is shown and colored purple.


Fig. 4. Ribbon diagrams of 3D models of Hyal-2, Hyal-3, Hyal-4, and HPH-20. All four models were obtained as described in the Figure 3 legend and in Materials and Methods. Similar to Hyal-1, the main, catalytic domain for each Hyal has the characteristic distorted $(\beta/\alpha)_8$ TIM barrel fold. The C-terminal domains are different for each Hyal. The molecules in all panels are shown in a similar orientation and use the same color scheme as that of Hyal-1 in Figure 3(A and B). The catalytic and C-terminal domains, as well as the N- and C-termini are shown and labeled. (**A**) Model of Hyal-2 degrading hyaluronan in human tissues to an approximate size of 50 disaccharides. (**B**) Structural model of human PH-20/SPAM1 hyaluronidase found in sperm and involved in fertilization, as well as facilitating penetration of sperm through the cumulus mass to the ovum. (**C**) Model of Hyal-3 enzyme. (**D**) Model of Hyal-4 with the specific chondroitinase activity.

**TABLE I. Domain Arrangement of Human Hyaluronidases[†]**

| Hyal | Hyal aa span | Domain | Domain aa span | Parent | Parent aa span | Functional annotations |
|---|---|---|---|---|---|---|
| Hyal-1 | 1–435 | 1 catalytic | 1–371 | 1FCQ[b] | 33–360 | hydrolase/hyaluronidase |
| | | 2 C-terminal | | NA[a] | NA[a] | unknown[d] |
| Hyal-2 | 1–471 | 1 catalytic | 1–376 | 1FCQ[b] | 33–366 | hydrolase/hyaluronidase |
| | | 2 C-terminal | 377–471 | NA[a] | NA[a] | unknown[d] |
| Hyal-3 | 1–463 | 1 catalytic | 1–373 | 1FCQ[b] | 41–366 | hydrolase/hyaluronidase |
| | | 2 C-terminal | 374–463 | NA[a] | NA[a] | unknown[d] |
| Hyal-4 | 1–481 | 1 catalytic | 1–384 | 1FCQ[b] | 41–366 | hydrolase/hyaluronidase |
| | | 2 C-terminal | 385–481 | NA[a] | NA[a] | unknown[d] |
| HPH-20 | 1–510 | 1 catalytic | 1–389 | 1FCQ[b] | 41–366 | hydrolase/hyaluronidase |
| | | 2 C-terminal | 390–510 | NA[a] | NA[a] | unknown[d] |
| BVHyal | 1–382 | 1 catalytic | 1–382 | 1FCQ[b] | 33–382[c] | hydrolase/hyaluronidase |

[†]The domains were predicted by the Ginzu domain parsing and fold detection method[15] and confirmed unanimously with very high confidence by all methods implemented in the META server/3D Jury method.[13,26] All Hyals have a two-domain architecture with the major domain being the catalytic one, with a very high primary, secondary, structural, and functional homology to the BVHyal of known 3D structure (PDB code 1FCQ). The structure of BVHyal consists of aa residues 33–382 of the full-length mature protein of this organism. The second and minor domain is located at the extreme C-terminus and has no homology to proteins with known structures. The primary structure conservation among all Hyals is significantly lower than for the catalytic domain. Their functions are still unknown.

[a]NA-Not available or unknown.

[b]PDB code is reported (3D structure coordinates available at *www.rcsb.org/pdb* under this code).

[c]The crystallized BVHyal enzyme consists of aa residues 33–382 of the full-length protein in this organism. Residues 33–43, 97–103, and 363–382 were not identified in the 1FCQ X-ray crystal structure.[6]

[d]The state-of-the-art sequence comparisons and structural analyses of the *ab initio* 3D model have not revealed any protein with similar sequence or 3D structure to the C-terminal region, domain, of the Hyals analyzed here.

catalytic part. These five models, therefore, represented reliable model structures for all five Hyal enzymes.

For all Hyals, two domain structures were predicted for each enzyme. The major, catalytic domain comprised aa residues starting from the very N-terminus of each protein (Fig. 2), and was always followed by a significantly smaller C-terminal domain. The catalytic domain model was constructed by its homology to BVHyal enzyme structure. The extreme N-terminal segments for the catalytic domain, ranging in length from 28 aa for Hyal-1 to 41 aa for HPH-20, were not included in the BVHyal template structure (and were thus omitted in the crystallized protein) (Fig. 2). This segment's structure was predicted, therefore, by the *ab initio* approach. Due to the short length of this N-terminal segment, the modeling in this part of Hyal is likely to be reliable. The sequence similarity among this portion of the protein was low compared to the remaining portion.

For the C-terminal domain, searches in structural databases showed no significant similarity to any known 3D structures. As a consequence, fold recognition and separate structural modeling by the Robetta server indicated a lack of potential independent homology modeling for this portion of the Hyal molecule. The structures of this domain were therefore constructed by the *ab initio* approach, as implemented by the Robetta server.[14,15] Even though the prediction methods have significantly improved recently, such *ab initio* models must still be examined with a high degree of caution. This C-terminal portion of the Hyal molecule was significantly smaller than the catalytic one, but larger than the N-terminal segment of the catalytic domain, ranging in size from 68 aa for Hyal-1 to 122 aa for HPH-20 (Table I, Fig. 2). The sequence similarity within this portion was minimal compared to the remaining part of the enzymes, including the extreme N-terminus. In summary, the entire catalytic domain structure was formulated based on the presence of the BVHyal homology model that included the major part of the catalytic domain, but that did not include the smaller N-terminal part. The extreme N-terminus and the C-terminal domain were predicted using the *ab initio* approach.

### Description of Model Structures

For all five Hyal structure models obtained, the major catalytic domain of the final models (Figs. 3 and 4) closely resembles the $(\beta/\alpha)_8$ TIM barrel of the parent BVHyal structure. The number of sheets constituting the central part of the TIM barrel varies, however, similar to that of various forms of the BVHyal structures available (PDB codes 1FCV, 1FCQ, and 1FCU). The wider end of the TIM barrel assumes a conformation that creates a large and elongated cleft. Its sides are built primarily from aa residues of loops formed between sheets and helices of the TIM barrel. This region was shown to be responsible for both binding the HA substrate and catalysis,[6] and is large enough to bind at least an octasaccharide segment of HA. The extreme N-terminal segment of the catalytic domain, predicted by the *ab initio* approach (see above), always assumes a helical conformation. It takes on a structure of a single helix for Hyal-1 and -2 (Hyals with the shortest segment at their N-terminus) or a set of two helices connected by a short loop for Hyal-3, Hyal-4, and HPH-20. Hyal-4 and HPH-20 are Hyals with longer aa segments at the N-terminus (Fig. 2). In all five cases, however, the helices are located on the outside of the TIM barrel core structure, are not incorporated in the distorted TIM barrel fold, and are primarily associated with the TIM barrel domain only. The interactions with the C-terminal domain

are minimal, if they exist at all (Figs. 3 and 4). The sequence similarity in this region of all Hyals is very low compared to the remaining part of the catalytic domain, and therefore it allows for slight variations in this segment of the Hyal structures.

For all Hyals at the C-terminus, the catalytic domain is followed by the second domain (major catalytic domain – domain 1; C-terminal short domain – domain 2), which is not present in the BVHyal structure template. The predicted structure of this domain varies significantly among different Hyals. It assumes a three-stranded antiparallel β-sheet, flanked on one side by a helix and on the other by low-complexity structures (extreme C-terminus) for Hyal-1. In Hyal-2 it is composed of a two-helix group, whereas in HPH-20 it is composed of a group of eight helices, three of which are very short. For Hyal-3 it assumes a two-helix structure and a significant region of low complexity. In Hyal-4, it is a group of three helices flanked on one side by a two-stranded antiparallel β-sheet (Figs. 3 and 4).

There is one commonality among all C-terminal Hyal structures, however. The C-terminal domain is always separated from the catalytic part by a linker peptide, presumably flexible and of significant length, making this domain relatively independent of the rest of the protein. This domain is likely to have its own, separate motion/dynamic properties and its own unique function, which is different for each Hyal (Figs. 3 and 4).

### Structure of Active Site

The active site was identified based on the homology to BVHyal[6] and earlier mutational analysis.[35] The active site is located within the substrate-binding cleft that transverses the wider end of the TIM barrel structure of the catalytic domain. The cleft does not extend to the C-terminal separate domain. The cleft is lined by a number of electropositive and hydrophobic aa residues, many of them conserved in all human Hyals. As such, it is ideally suited to bind the electronegative and hydrophobic HA/Ch/ChS substrate. Based on earlier investigations,[6] the active site is composed of a single catalytic residue, Glu131, and a group of several residues that position the carbonyl nucleophile of the HA substrate for catalysis. Structural alignment of all Hyals, in addition to the 3D alignment of all modeled Hyal structures presented here and that of the BVHyal X-ray structure, the BVHyal X-ray structure complex with the HA substrate, and the BPH-20 enzyme modeled elsewhere demonstrated structural conservation (defined as the same positioning) of this catalytic Glu131 (Hyal-1 numbering) residue and its accompanying group of substrate-positioning residues, Asp129, Tyr202, Tyr247, and Trp321 (Hyal-1 numbering) (Fig. 5). All of these residues are also strictly conserved in sequence (Table II, Figs. 2 and 5, and data not shown). The Tyr247 aa residue is replaced by Cys263 in the Hyal-4 enzyme. This Cys263 is one of the positioning residues of the nucleophile/base of the substrate and may reflect this enzyme's specificity for chondroitin (Table II). Due to such sequence and structural conservation of the residues involved in catalysis among human Hyals and with the BVHyal enzyme, it was reasonable to assume that the mechanism of all human Hyals is the same as that proposed for BVHyal.

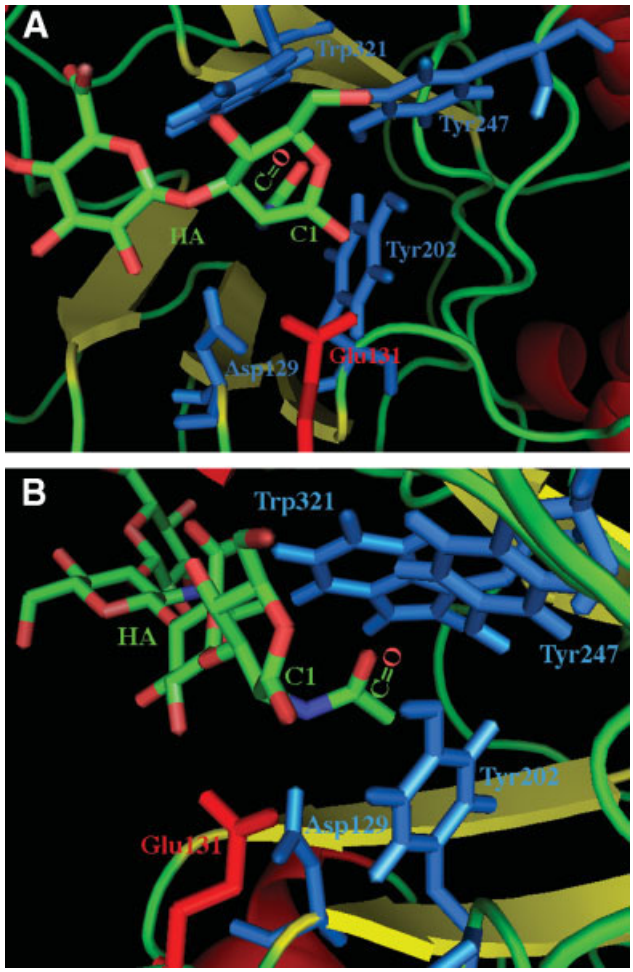


Fig. 5. Catalytic part of the Hyal-1 HA-binding cleft. The scheme is based on the structure of the Hyal-1 enzyme with an HA tetrasaccharide positioned as in the structure of the BVHyal homolog (pad coordinate code: 1FCV). The reducing end of the bound HA molecule is located in the HA-binding cleft and is depicted in ball-and-stick fashion colored in by atom type (C, green; N, blue; O, red; S, yellow). The catalytic residue, Glu131, is also shown in ball-and-stick fashion colored in red (labeled); other residues that position the carbonyl of the acetylamido group (labeled) of HA are also shown and are colored in blue (labeled). The C1 atom and carbonyl group (marked as C═O) of HA's *N*-acetyl-*D*-glucosamine are labeled. (**A**) View from above HA and HA-binding cleft towards the enzyme's catalytic domain. The catalytic H-donor (acid) Glu131, in red, is shown in position to interact with the C1 carbon. The carbonyl residues, in blue, are deeper in the enzyme's cleft where the acetamido group is located. (**B**) View from down the HA chain along the HA-binding cleft. The carbonyl positioning residues are clearly in position to modify the position of the acetamido group, to allow interaction of the carbonyl group with the C1 carbon.

Essential to the catalytic reaction is Glu113, proposed to act as an acid, an H-donor. The substrate-positioning residues are involved in a precise orientation of the carbonyl of the acetamido group of HA. This carbonyl group of the substrate serves as a nucleophile instead of the second acidic residue that would normally perform this

**TABLE II. Numbering Scheme for Conserved Residues Involved in Catalytic Process and in Essential Positioning of Substrate's Carbonyl of the Acetamido Group**

| BVHyal | BPH-20 | Hyal-1 | Hyal-2 | Hyal-3 | Hyal-4 | HPH-20 |
|---|---|---|---|---|---|---|
| Catalytic residue: | | | | | | |
| Glu113 | 149 | 131 | 135 | 129 | 147 | 148 |
| Positioning residues: | | | | | | |
| Asp111 | 147 | 129 | 133 | 127 | 145 | 146 |
| Tyr184 | 220 | 202 | 206 | 202 | 218 | 219 |
| Tyr227 | 265 | 247 | 253 | 246 | Cys263 | 264 |
| Trp301 | 341 | 321 | 327 | 319 | 339 | 339 |

The residues were divided into the catalytic Glu residue and the supporting residues that position the HA's carbonyl of the acetamido group for catalysis. The Cys263 residue of Hyal-4 breaks the conserved scheme and likely reflects this Hyal's specificity for chondroitin and its chondroitinase function. All other residues are strictly conserved (Fig. 2).

function. In this case, there is no enzyme–substrate covalent intermediate formed, as would occur in the usual situation.

### Scheme for Hyal-Catalyzed Degradation of HA

Due to high structural and sequence homology, it is assumed that human Hyals degrade HA by the same mechanism as the BVHyal enzyme.[6] The degradation of HA by human Hyal hydrolases proceeds, therefore, *via* the double-displacement mechanism with retention of the HA substrate conformation, and involves one glutamic acid aa residue as an H-donor (Glu131 in Hyal-1) and a carbonyl oxygen of the N-acetyl group of HA performing the function usually assigned to another carboxylic aa. The regular, two-carboxylic-acid mechanism can be modified to reflect the different nucleophilic residue. Therefore, the steps involved in this mechanism are as follows (Figs. 5 and 6):

1. Binding of Hyal to the polysaccharide HA substrate.
2. Cleft residues surrounding the catalytic site positioning the carbonyl oxygen nucleophile of the HA's N-acetyl group next to the to-be-cleaved β1,4 glycosidic bond, to attack the C1 carbon of the same sugar to form a covalent intermediate between them. This leads to cleavage of the glycosidic bond in the substrate on the non-reducing side of the glycosidic oxygen. The process also results in the inversion of the anomeric C1 atom configuration.
3. At the same time, the protonated Glu131 donates its H (deprotonation, acid function) to the glycosidic oxygen, leaving part of HA (glycan on the reducing side of the glycosidic bond) cleaved.
4. Hydrolytic cleavage of the carbonyl oxygen–C1 intermediate bond by a water molecule in the active site leading to re-protonation of Glu131, readying it for the next catalytic step, and leading to the second inversion of the configuration of C1.
5. Release of the remaining HA product from the Hyal's active site (glycan on the non-reducing side of the cleaved glycosidic bond).

The anomeric configuration of the C1 carbon atom of the substrate is retained after its conformation is inverted twice during the catalysis (steps 2 and 4 above). The formation of an oxocarbonium-ion transition state has been implicated in this process at step 2 described above. Structural evidence from the BVHyal–HA tetrasaccharide complex suggested that (i) as the carbonyl nucleophile moves into place to interact with the C1 carbon of HA, it results in (ii) the change of puckering of the pyranose ring of *N*-acetyl-*D*-glucosamine on the non-reducing side of the bond to be cleaved, from a regular chair to a distorted boat and (iii) as a consequence, by this distortion of the chair conformation to a near coplanar arrangement of this part of HA (a near coplanar arrangement of C2-C3-C5-Oρ-C1-$O_\gamma$) the glycosidic bond is displaced, which results in (iv) positioning this bond close to the Glu131 (Hyal-1 numbering) to allow for the donation of its H to this glycosidic O as the bond is being cleaved.[6]

The site-directed mutagenesis in human sperm PH-20 enzyme of Glu148 and Asp146 (equivalent of Glu131 and Asp129 in Hyal-1) to Gln and Asn, respectively, results in the total loss of activity for E148Q mutant and only 3% activity for D146N construct.[36] These results are entirely consistent with the proposed mechanism, and support the catalytic importance of this single Glu aa residue. The residual low level of activity of the D146N mutant in HPH-20 is also consistent with it having only a supportive role, being only one positioning residue out of a total of four (Fig. 5, Table II).

### Endolytic, Random Cut Pattern of Enzyme Action

The process of HA degradation proceeds through a random bite type of catalysis, an endolytic, non-processive mechanism, generating products of degradation that are substrates for further cleavage. This leaves the active site of Hyals available after each cleavage reaction. The shortest HA fragment that can still be cleaved by mammalian enzymes is a hexasaccharide unit.[3,37] This hexasaccharide also competes successfully in several important biological processes; for example, it binds to CD44, competing with long HA chains.[38] As a consequence, the final product of HA digestion by human Hyals is an HA tetrasaccharide.

Mammalian Hyals also degrade chondroitins, albeit at a significantly slower rate. The mechanism described above should be the same for this substrate. The sulfation pattern of chondroitin sulfates can presumably be accom-

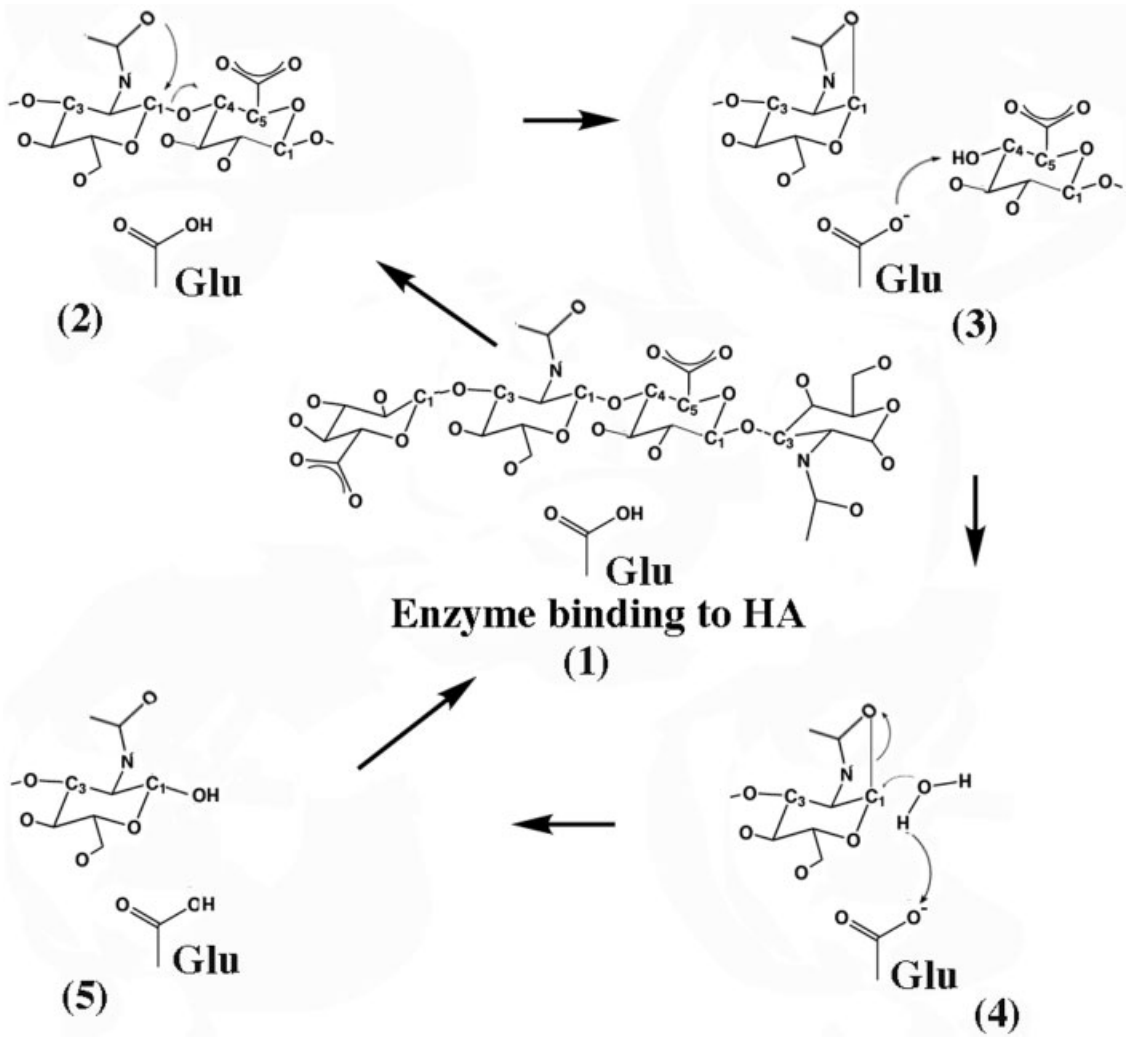


Fig. 6. Schematic double-displacement retaining mechanism, common to all human Hyals. This mechanism is characteristic of polysaccharide hydrolases and involves the five depicted steps, 1–5. It involves a double-displacement mechanism with the retention of C1 carbon configuration (retaining enzyme). The catalytic process proceeds through an intermediate reaction step involving the oxocarbonium ion at the C1 position of *N*-acetyl-*D*-glucosamine of the HA substrate. The catalytic acid function is performed by a Glu residue (absolutely conserved among all Hyals), and the nucleophile/base function is attributed to the carbonyl oxygen of C-2 acetamido group of the substrate: HA, Ch, or ChS.

modated by these Hyal enzymes, and allows for binding and catalysis. This is not entirely unexpected, as the notion that Hyals evolved from chondroitinases appears to be well supported. Chondroitin appears earlier in evolution than does HA.[3,38,39] Hyals may have evolved originally from preexisting chondroitinases, retaining residual chondroitinase activity and recognizing the ancestral substrate, while becoming specialized for HA degradation.

### Putative Functions of C-Terminal Domains

The C-terminal domains have the greatest variability, both sequence and structural, among all Hyal enzymes. The additional functions attributed to these multifunctional enzymes may well lie within these domains. The lack of experimental data and the lack of sequence/structural homologs for the C-terminal domains make functional studies difficult and speculative at best. However, the presence of a peptide linker between the catalytic domain of Hyals and the C-terminal domain is indicative of structural independence, and of dynamic flexibility between this domain and the remainder of the enzyme. Consistent with this is the significant length and unstructured nature of the linker. Such an arrangement would also be indicative of an independent, supportive role of the C-terminal domain in the catalytic function of the entire enzyme. The different structures for this domain among all five Hyals are also indicative of different functions or tissue specificity for this domain, in the same way that each of the Hyals has a unique pattern of tissue expression. Analysis of the structural properties of this domain and its surface indicates HA-binding ability.

All Hyal structural models of the C-terminal domain have large grooves/clefts that can accommodate HA substrate binding. There is no experimental evidence, however, for this. Without such evidence, explanations also

cannot be posited for differences in functional properties of the various Hyals, including an ability to account for the different-sized degradation products generated by Hyal-1 and -2.

There are a number of speculative possibilities for the function of the C-terminal domain. This domain may be involved in interactions with the HA substrate, and may (i) increase the affinity of the Hyals for an HA-rich matrix, and therefore (ii) facilitate co-localization of enzyme and substrate. A further possible consequence of binding to HA might be (iii) the disruption of non-covalent interactions between substrate chains or between substrate and other polysaccharides.[40] Also, it is possible that the domain plays a role in (iv) the appropriate orientation of the substrate with respect to the catalytic domain, which would further facilitate HA degradation in a specific orientation of HA chains with respect to enzyme, as by reducing end-directional positioning. Such directional positioning of HA is observed for the bacterial Hyal enzymes. These have been studied in greater structural detail than have the human Hyals.[9,41,42] All of these reasons would facilitate HA cleavage. Flexibility and solvent exposure of the peptide linker also suggests that this site would be the initial site for autodegradation and/or degradation by external proteases, as has been observed for the bacterial Hyals[43–45] and BPH-20. The deletion of 36 residues at the C-terminus of HPH-20 resulted in a decrease of this enzymes activity to 25%, but the truncated enzyme was not released from the expressing cells, making its activity determination uncertain. A similar polypeptide of PH-20 from the cynomolgus monkey was secreted by expressing cells, purified, and shown to retain full activity.[46] The catalytic independence of the PH-20 enzyme's C-terminus, suggested by our modeling, was confirmed by the preservation of full activity by such truncates.

The bioinformatics/structural results presented here demonstrate the need for further research in determining the function of the C-terminal domain, and delineate possible avenues for such studies. It is not coincidental that the bacterial Hyals are also multidomain enzymes. In their case, however, it is the extreme N-terminal domain that is implicated in HA binding, as it parallels the C-terminal processes described above.[42]

## CONCLUSIONS

The present studies propose 3D model structures for the human Hyal enzymes: Hyal-1–4 and PH-20. The models indicate that very similar structures exist and differ primarily in the structure of their C-terminal domains. These are of unknown function. However, the catalytic cleft and the active sites are highly conserved. The models allow for clear identification of the catalytic machinery used by Hyals. Such machinery involves an acidic residue, Glu131 (Hyal-1 numbering) functioning as an H-donor. There are no other acidic residue, such as Glu or Asp, in the active sited that could support such a function. Therefore, as is the case for some chitinases, these enzymes utilize the substrate carbonyl group of the acetamido moiety as a nucleophile. No enzyme–substrate covalent intermediate is created in this process, in marked contrast to other glycan-degrading hydrolases.

The sequence analysis, fold recognition, and modeling studies allow for identification of the residues involved in Hyal catalysis and are of significant utility. The model structures of the C-terminal domain in particular must be viewed with caution. Further experimental evidence is essential in delineating the function of this part of the Hyal models. Additional data must be obtained and a larger database of 3D structures generated with annotated functions for the various domains to further enhance our understanding of this class of enzymes, which until recently had been relatively neglected.

## ACKNOWLEDGMENTS

This work was supported by a grant from the National Institutes of Health and by the California Spinal Cord Injury Research Fund.

## REFERENCES

1. Coutinho PM, Henrissat B. Carbohydrate-active enzymes server http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html. 1999.
2. Coutinho PM, Henrissat B. Carbohydrate-active enzymes: an integrated database approach. In: Gilbert HJ, Davies G, Henrissat B, Svensson B, editors. Recent advances in carbohydrate bioengineering. Cambridge, UK: The Royal Society of Chemistry; 1999. p 3–12.
3. Stern R. Devising pathway for hyaluronan catabolism. Are we there yet? Glycobiology 2003;13(12):105R–115R.
4. Heckel D, Comtesse N, Brass N, Blin N, Zang KD, Meese E. Novel immunogenic antigen homologous to hyaluronidase in meningioma. Hum Mol Genet 1998;7(12):1859–1872.
5. Koshland DE. Sterochemistry and the mechanism of enzymatic reactions. Biol Rev Camb Philos 1953;28:416–436.
6. Markovic-Housley Z, Miglierini G, Soldatova L, Rizkallah PJ, Muller U, Schirmer T. Crystal structure of hyaluronidase, a major allergen of bee venom. Structure Fold Des 2000;8(10):1025–1035.
7. Davies GJ, Dauter M, Brzozowski AM, Bjornvad ME, Andersen KV, Schulein M. Structure of the *Bacillus agaradherans* family 5 endoglucanase at 1.6 Å and its cellobiose complex at 2.0 Å resolution. Biochemistry 1998;37(7):1926–1932.
8. Divne C, Stahlberg J, Teeri TT, Jones TA. High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. J Mol Biol 1998;275(2):309–325.
9. Jedrzejas MJ. Structural and functional comparison of polysaccharide-degrading enzymes. Crit Rev Biochem Mol Biol 2000;35(3): 221–251.
10. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22(22):4673–4680.
11. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 1999;292(2):195–202.
12. Bujnicki JM, Elofsson A, Fischer D, Rychlewski L. Structure prediction meta server. Bioinformatics 2001;17(8):750-751.
13. Ginalski K, Elofsson A, Fischer D, Rychlewski L. 3D-Jury: a simple approach to improve protein structure predictions. Bioinformatics 2003;19(8):1015–1018.
14. Chivian D, Kim DE, Malmstrom L, Bradley P, Robertson T, Murphy P, Strauss CE, Bonneau R, Rohl CA, Baker D. Automated prediction of CASP-5 structures using the Robetta server. Proteins 2003;53 Suppl 6:524–533.
15. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 2004;32(Web Server issue):W526–531.
16. Sippl MJ. Recognition of errors in three-dimensional structures of proteins. Proteins 1993;17(4):355–362.
17. Eisenberg D, Luthy R, Bowie JU. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol 1997;277:396–404.

18. Wallner B, Elofsson A. Can correct protein models be identified? Protein Sci 2003;12(5):1073–1086.
19. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr A 1991;47(Pt 2):110–119.
20. DeLano WL. The PyMOL Molecular Graphics System: www.pymol.org; 2003.
21. Carson M. Ribbons. Methods Enzymol 1997;277:493–505.
22. Nicholls A, Sharp KA, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11(4):281–296.
23. Botzki A, Rigden DJ, Braun S, Nukui M, Salmen S, Hoechstetter J, Bernhardt B, Dove S, Jedrzejas MJ, Buschauer A. L-ascorbic acid-6-hexadecanoate, a potent hyaluronidase inhibitor: X-ray structure and molecular modeling of enzyme-inhibitor complexes. J Biol Chem 2004;279(44):45990–45997.
24. Holm L, Sander C. Touring protein fold space with Dali/FSSP. Nucleic Acids Res 1998;26(1):316–319.
25. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22(22):4673–4680.
26. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997; 25(17):3389–3402.
27. Lundstrom J, Rychlewski L, Bujnicki J, Elofsson A. Pcons: a neural-network-based consensus predictor that improves fold recognition. Protein Sci 2001;10(11):2354–2362.
28. Ginalski K, Rychlewski L. Detection of reliable and unexpected protein fold predictions using 3D-Jury. Nucleic Acids Res 2003; 31(13):3291-3292.
29. Andreeva A, Howorth D, Brenner SE, Hubbard TJ, Chothia C, Murzin AG. SCOP database in 2004: refinements integrate structure and sequence family data. Nucleic Acids Res 2004;32 Database issue:D226–229.
30. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 1995;247(4):536–540.
31. Aleshin A, Golubev A, Firsov LM, Honzatko RB. Crystal structure of glucoamylase from *Aspergillus awamori* var. X100 to 2.2-Å resolution. J Biol Chem 1992;267(27):19291–19298.
32. Perrakis A, Tews I, Dauter Z, Oppenheim AB, Chet I, Wilson KS, Vorgias CE. Crystal structure of a bacterial chitinase at 2.3 Å resolution. Structure 1994;2(12):1169–1180.
33. Rigden DJ, Bagyan I, Lamani E, Setlow P, Jedrzejas MJ. A cofactor-dependent phosphoglycerate mutase homolog from *Bacillus stearothermophilus* is actually a broad specificity phosphatase. Protein Sci 2001;10(9):1835–1846.
34. Rigden DJ, Setlow P, Setlow B, Bagyan I, Stein RA, Jedrzejas MJ. PrfA protein of *Bacillus* species: prediction and demonstration of endonuclease activity on DNA. Protein Sci 2002;11(10):2370–2381.
35. Arming S, Strobl B, Wechselberger C, Kreil G. *In vitro* mutagenesis of PH-20 hyaluronidase from human sperm. Eur J Biochem 1997;247(3):810–814.
36. Stern R. The mammalian hyaluronidases. Update. In: Hascall VC, Yanagishita M, editors. Science of hayluronan today: Glycoforum www.glycoforum.gr.jp/science/hyaluronan; 2004.
37. Ghatak S, Misra S, Toole BP. Hyaluronan oligosaccharides inhibit anchorage-independent growth of tumor cells by suppressing the phosphoinositide 3-kinase/Akt cell survival pathway. J Biol Chem 2002;277(41):38013–38020.
38. DeAngelis PL. Evolution of glycosaminoglycans and their glycosyltransferases: implications for the extracellular matrices of animals and the capsules of pathogenic bacteria. Anat Rec 2002;268(3): 317–326.
39. Rigden DJ, Jedrzejas MJ. Structures of *Streptococcus pneumoniae* hyaluronate lyase in complex with chondroitin and chondroitin sulfate disaccharides: insights into specificity and mechanism of action. J Biol Chem 2003;278(50):50596–50606.
40. Tomme P, Warren RA, Miller RC, Jr., Kilburn DG, Gilkes NR. Cellulose-binding domains: classification and properties. In: Sadler JN, Penner M, editors. Enzymatic degradation of insoluble polysaccharides. Washington, D.C.: American Chemical Society; 1995. p 142–146.
41. Jedrzejas MJ, Mello LV, de Groot BL, Li S. Mechanism of hyaluronan degradation by *Streptococcus pneumoniae* hyaluronate lyase. Structures of complexes with the substrate. J Biol Chem 2002;277(31):28287–28297.
42. Rigden DJ, Jedrzejas MJ. Genome-based identification of a carbohydrate binding module in *Streptococcus pneumoniae* hyaluronate lyase. Proteins 2003;52(2):203–211.
43. Jedrzejas MJ, Chantalat L, Mewbourne RB. Crystallization and preliminary X-ray analysis of *Streptococcus pneumoniae* hyaluronate lyase. J Struct Biol 1998;121(1):73–75.
44. Jedrzejas MJ, Mewbourne RB, Chantalat L, McPherson DT. Expression and purification of *Streptococcus pneumoniae* hyaluronate lyase from *Escherichia coli*. Protein Expr Purif 1998;13(1):83–89.
45. Jedrzejas MJ, Chantalat L. Structural studies of *Streptococcus agalactiae* hyaluronate lyase. Acta Crystallogr D Biol Crystallogr 2000;56(Pt 4):460–463.
46. Lin Y, Mahan K, Lathrop WF, Myles DG, Primakoff P. A hyaluronidase activity of the sperm plasma membrane protein PH-20 enables sperm to penetrate the cumulus cell layer surrounding the egg. J Cell Biol 1994;125(5):1157–1163.