

# Alternate Structural Conformations of *Streptococcus pneumoniae* Hyaluronan Lyase: Insights into Enzyme Flexibility and Underlying Molecular Mechanism of Action

Daniel J. Rigden<sup>1,2</sup>, James E. Littlejohn<sup>1</sup>, Harshad V. Joshi<sup>3</sup>  
Bert L. de Groot<sup>3</sup> and Mark J. Jedrzejewski<sup>1\*</sup>

<sup>1</sup>Children's Hospital Oakland  
Research Institute, Oakland  
CA 94609, USA

<sup>2</sup>School of Biological Sciences  
University of Liverpool  
Liverpool L69 7ZB, UK

<sup>3</sup>Computational Biomolecular  
Dynamics Group, Max Planck  
Institute for Biophysical  
Chemistry, Am Fassberg 11  
37077 Göttingen, Germany

*Streptococcus pneumoniae* hyaluronan lyase is a surface enzyme of this Gram-positive bacterium. The enzyme degrades several biologically important, information-rich linear polymeric glycans: hyaluronan, unsulfated chondroitin, and some chondroitin sulfates. This degradation facilitates spreading of bacteria throughout the host tissues and presumably provides energy and a carbon source for pneumococcal cells. Its  $\beta$ -elimination catalytic mechanism is an acid/base process termed proton acceptance and donation leading to cleavage of  $\beta$ -1,4 linkages of the substrates. The degradation of hyaluronan occurs in two stages, initial endolytic cuts are followed by processive exolytic cleavage of one disaccharide at a time. In contrast, the degradation of chondroitins is purely endolytic.

Structural studies together with flexibility analyses of two streptococcal enzymes, from *S. pneumoniae* and *Streptococcus agalactiae*, allowed for insights into this enzyme's molecular mechanism. Here, two new X-ray crystal structures of the pneumococcal enzyme in novel conformations are reported. These new conformations, complemented by molecular dynamics simulation results, directly confirm the predicted domain motions presumed to facilitate the processive degradative process. One of these new structures resembles the *S. agalactiae* enzyme conformation, and provides evidence of a uniform mechanistic/dynamic behavior of this protein across different bacteria.

© 2006 Elsevier Ltd. All rights reserved.

**Keywords:** bacteria; chondroitin; glycan degradation; molecular mechanism; molecular dynamics

\*Corresponding author

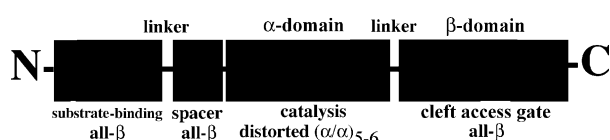
## Introduction

Hyaluronan lyase (Hyal) is one of the major surface proteins of many *Streptococcus* species and

has also been found in species such as *Enterococcus*, *Staphylococcus*, *Streptomyces*, and *Propionibacterium*. In *Streptococcus pneumoniae* Hyal is present on the majority of strains of this organism examined to date and its expression is directly related to virulence of this organism.<sup>1,2</sup> The enzyme degrades essential components of the host's extracellular matrix (ECM), hyaluronan (HA), unsulfated chondroitin (CH), and certain chondroitin sulfates (CHS), and as such it facilitates bacterial spread among the host tissues. By doing so, it also provides an abundant source of carbon and energy for bacterial cells. The enzyme is found in two forms, the first closely associated with bacterial cells, the second a cell-free form presumably released from bacterial cells by the action of sortase enzymes.<sup>2-4</sup>

Abbreviations used: CH, unsulfated chondroitin; CHS, chondroitin sulfate; ECM, extracellular matrix of tissues; HA, polymeric hyaluronan; Hyal, hyaluronan lyase; OPLS-AA, optimized potentials for liquid simulations—all atom; PAD, proton acceptance and donation; PDB, Protein Data Bank; PCA, principal component analysis; PME, particle mesh Ewald; rmsd, root-mean-square deviation; SagHyal, *S. agalactiae* Hyal; SpnHyal, *S. pneumoniae* Hyal; MD, molecular dynamics.

E-mail address of the corresponding author:  
mjedrzejewski@chori.org



**Figure 1.** Modular structure of bacterial Hyals. Bacterial Hyals are composed of four domains, a substrate-binding module, a spacer domain, a major catalytic domain ( $\alpha$ -domain), followed by a C-terminal domain ( $\beta$ -domain) that regulates HA access to the cleft.<sup>6</sup> The N-terminal substrate-binding module and the C-terminal domain are connected to the rest of the protein by peptide linkers.

The cell-attached Hyal is anchored to the surface of bacterial cell wall through a covalent linkage to the cross-bridges of the wall's peptidoglycan.<sup>5,6</sup> As is the case for many glycan active enzymes, Hyal molecules contain various distinct domains, each with a well-defined and specific function related to the enzymatic role of this protein. Hyal is a four-domain protein consisting of a presumed N-terminal HA/CH/CHS-binding domain, followed by a spacer domain, the catalytic domain (termed  $\alpha$ -domain), and the C-terminal domain (termed  $\beta$ -domain) that modulates the access of the substrate to the catalytic machinery (Figure 1).<sup>6-8</sup> The full-length four-domain pneumococcal molecule in its recombinant form from *Escherichia coli* is unstable and undergoes degradation by a (auto)-proteolytic process.<sup>9,10</sup> The largest stable enzyme is composed of the two C-terminal domains,  $\alpha$  and  $\beta$ -domain, and retains essentially full catalytic activity *in vitro*.<sup>2,9</sup> The structure of these two major domains has been elucidated by X-ray crystallography<sup>8</sup> and was followed by a variety of structural,<sup>11-13</sup> biochemical and molecular biology studies.<sup>8,12,14</sup> These studies have led to the formulation of the proposed mechanism of catalysis for this enzyme<sup>8</sup> and its mode of action during degradation of different substrates.<sup>12-14</sup> The full-length enzyme contains additional residues at the N terminus arranged in two domains<sup>6</sup> that are not catalytically important (Figure 1). The first of these domains is clearly a substrate (HA/CH/CHS) binding domain that is thought to act as to enhance the Hyal affinity for substrates.<sup>6,15</sup> The second, small domain acts as a spacer to distance the

catalytic domain from the substrate-binding domain at the extreme N terminus and thus to facilitate processive mechanism of feeding the substrate for catalysis.<sup>6</sup> In addition, flexibility studies revealed several significant domain motions and allowed for their putative associations with specific roles necessary for catalysis and processivity.<sup>13</sup> One of these principal modes of domain motion was found to correspond to a structural difference evident from a comparison of the first pneumococcal Hyal structure (*SpnHyal*) with that of the *Streptococcus agalactiae* enzyme.<sup>7,16</sup> (*SagHyal*; see Table 1 for nomenclature of the main crystal structures discussed here). However, at that stage the possibility could not be excluded that such structural differences were due to sequence differences between the two bacterial Hyals.

The main feature of the catalytic  $\alpha$ -domain is the presence of a large, long cleft where the substrates bind and undergo catalysis. This domain exhibits high flexibility that presumably facilitates the mechanism of action of substrate degradation. A flexibility analysis identified three distinct domain motions involving:<sup>13,17</sup> (i) a rotation/twisting motion of the sides of the cleft that yields a  $\sim 10$  Å movement along the cleft axis; (ii) an opening/closing of the substrate-binding cleft accompanied by movement of catalytic amino acid residues and those responsible for the release of the product of degradation. The third movement (iii) involves a shift in position of the  $\beta$ -domain resulting in an effective opening and closing of the access/entrance to the cleft. This motion depicts the functional role of the  $\beta$ -domain as the modulator of access to the enzyme's cleft. In a putative model, these motions are directly related to the enzyme's function: catalysis, endolytic "random-bite" degradation, and exolytic cleavage of one disaccharide at a time from the reducing to non-reducing direction in a processive manner as has been described.<sup>18</sup> Of the simulated domain motions, the opening/closing mode of the cleft (ii) was found to agree closely with the observed differences between the crystal structures of *SagHyal* and *SpnHyal*.<sup>17</sup>

In the current work we report the structures of *S. pneumoniae* Hyal in two new conformations that directly reflect the flexibility and motion of the enzyme. These structures confirm the

**Table 1.** Crystal structures referred to in this work and their short names

Short name	Bacterium	Ligand	Resolution (Å)	PDB code	Reference
<i>SpnHyal</i>	<i>S. pneumoniae</i>	None	1.56	1egu	8
<i>SpnHyal-holo</i>		Hyaluronan hexasaccharide/ tetrasaccharide/disaccharide	2.00/1.53/1.70	1loh/1lxx/1c82	11,13
<i>SpnHyal-<i>apo</i></i>		None <sup>a</sup> /native	2.00/1.56	1loh <sup>a</sup> /1egu	13
<i>SpnHyal</i> -PEGMME		None	2.80	2brw	This work
<i>SpnHyal</i> -MALONATE		None	3.30	2brv	This work
<i>SpnHyal</i> -vitamin C		Vitamin C	2.00	1f9g	21
<i>SagHyal</i>	<i>S. agalactiae</i>	None	2.10	1f1s	7
<i>SagHyal-holo</i>		Hyaluronan hexasaccharide	2.20	1lxx	17

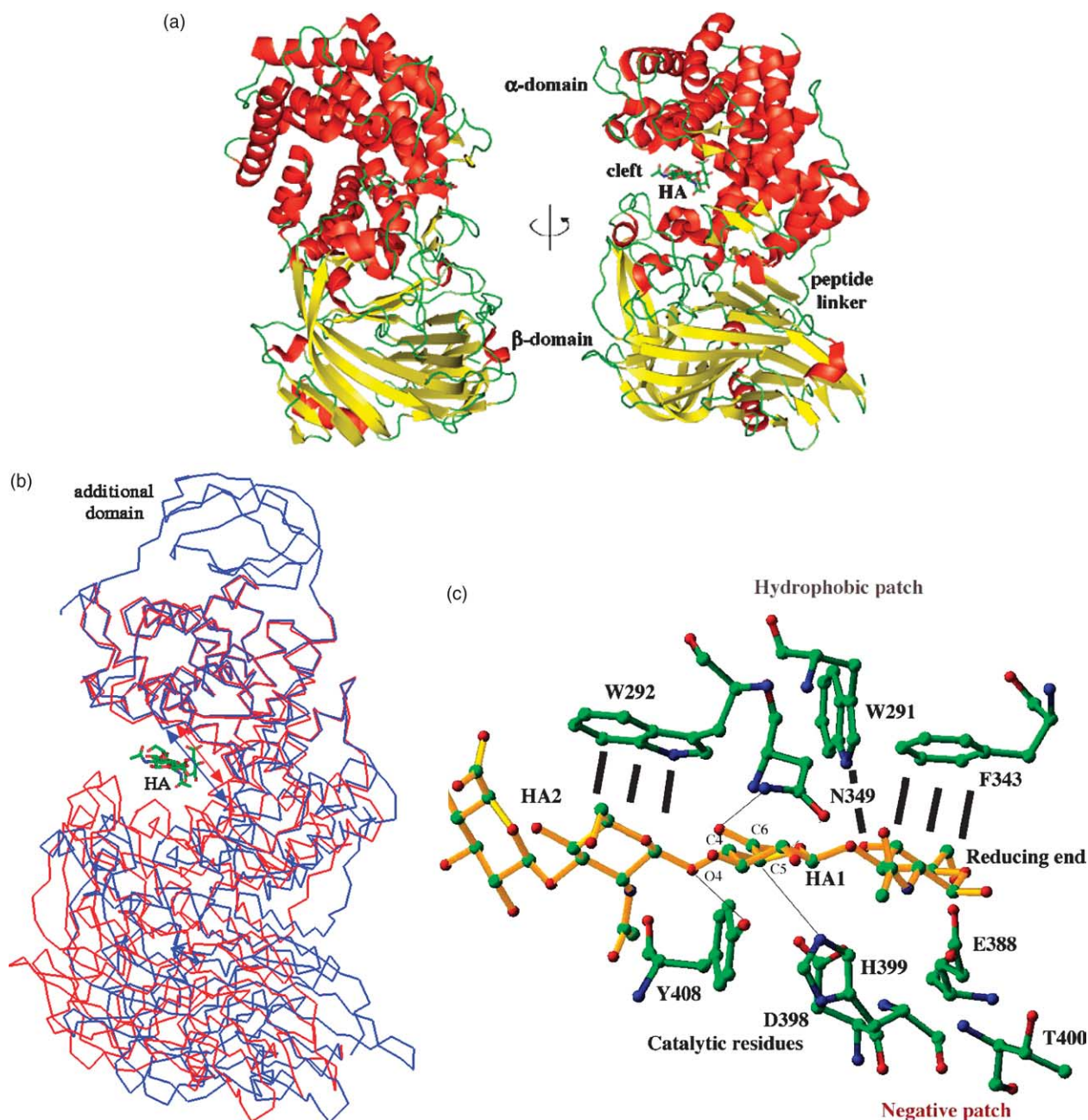
<sup>a</sup> The *apo* structure used for the MD studies is identical with the *holo* structure of the complex (PDB code 1loh) but with HA substrate computationally removed. The native structure of the *S. pneumoniae* enzyme was also reported (PDB code 1egu).<sup>8</sup>

previously predicted flexibility properties of this enzyme as well as its functional and mechanistic properties. The new structures also reveal that the structural differences previously observed for different bacterial Hyals are not primarily due to sequence differences. Rather, these structures, together with additional molecular dynamics simulation results, reflect a set of specific domain motions likely to be common to all bacterial Hyal enzymes.

## Results

### Description of the new pneumococcal Hyal structures

The individual domain structures of Hyal in the two new conformations, *SpnHyal*-PEGMME and *SpnHyal*-MALONATE, are similar to those observed in earlier crystal structures<sup>8,11,13</sup> (Figure 2;



**Figure 2.** Three-dimensional structures of (a) Hyal from *S. pneumoniae* in perpendicular views with bound hexasaccharide substrate, and (b) the catalytic portion of the pneumococcal enzyme cleft with the bound substrate and residues that are important for catalysis. The domains, the catalytic cleft, residues, and the substrate are labeled. The atoms are color-coded by atom type: green, carbon; red, oxygen; blue, nitrogen. HA substrate bonds are in yellow and protein residue bonds are in green. The hydrophobic patch positioning interactions with the HA substrate are depicted by thick black lines. HA1 and HA2 are HA disaccharides. The glycosidic bond between HA1 and HA2 is cleaved during the reaction. The thin lines represent hydrogen bonds or, in the case of H399, its relationship to C5 whose proton it is thought to abstract during the reaction.

**Table 2.** Crystal X-ray diffraction and structure refinement statistics

	<i>SpnHyal</i> - PEGMME	<i>SpnHyal</i> - MALONATE
Space group	P21	C2
Unit cell dimensions		
<i>a</i> (Å)	83.4	117.7
<i>b</i> (Å)	83.3	101.1
<i>c</i> (Å)	98.7	85.2
$\gamma$ (deg.)	98.8	125.1
Low resolution diffraction limit (Å)	95.4	69.0
High resolution diffraction limit (Å)	2.80	3.30
Completeness <sup>a</sup>	88.4 (78.2)	93.8 (84.0)
$I/\sigma(I)$ <sup>a</sup>	7.6 (2.2)	6.3 (2.3)
Multiplicity <sup>1a</sup> (%)	2.8 (2.6)	4.4 (5.0)
$R_{\text{merge}}$ <sup>a</sup> (%)	12 (69)	17 (61)
Non-hydrogen protein atoms	11582	5799
Sulfate atoms	10	0
Malonate atoms	0	42
Non-hydrogen solvent atoms	112	0
Number of reflections <sup>1</sup>	29,209 (4413)	11,046 (1526)
$R$ (%) <sup>a</sup>	23.1 (31.7)	30.2 (51.3)
$R_{\text{free}}$ (%) <sup>a</sup>	27.7 (38.4)	32.7 (52.3)
Mean temperature factor $B$ (Å <sup>2</sup> )		
All atoms	33.1	55.1
Protein	33.2	54.4
Protein main-chain	31.4	54.1
Protein side-chain	34.9	54.6
Sulfate or malonate	67.2	108.8
Solvent	20.7	–
rms deviation from ideal values		
Bond lengths (Å)	0.008	0.007
Bond angles (deg.)	1.59	1.08

<sup>a</sup> Values in parentheses are for the highest resolution shell (*SpnHyal*-PEGMME, 2.95–2.80 Å; *SpnHyal*-MALONATE, 3.48–3.30 Å).

**Table 2).** The nomenclature PEGMME and MALONATE indicates the main crystallization agent used to obtain diffracting crystals of this molecule in two new conformations. The crystallized Hyal molecule is composed of two domains, the catalytic domain ( $\alpha$ -domain) having an  $\alpha_5/\alpha_5$  barrel structure and the  $\beta$ -domain comprised mainly of an antiparallel, three-layer  $\beta$ -sandwich) (**Figure 2**). As before, the catalytic cleft transverses the  $\alpha$ -domain. The active site of the enzyme is located within the barrel domain on one side of the cleft area.<sup>13,19–21</sup> Both new structural conformations of the pneumococcal Hyal enzyme have structural differences from previously observed Hyal crystal structures. These were analyzed using intra-domain root-mean-square deviation (rmsd) calculations and Dyndom to describe domain motions. These differences are of quite different nature in the two new structures.

In *SpnHyal*-PEGMME there are two molecules in the asymmetric unit. Specific local conformational differences were observed and accounted for during refinement, but the two chains are nevertheless very similar; 721  $C^\alpha$  atoms may be fit with an rmsd of just 0.3 Å. For the purposes of other analyses, one chain (chain A) only was used.

The  $\beta$ -domain (residues 539–890) superimposes well on the corresponding domain of the original Hyal structure (*SpnHyal*; see **Table 1**), 352  $C^\alpha$  atoms superimpose with an rmsd of 0.26 Å. The  $\alpha$ -domain (residues 170–537) fits much less well yielding an rmsd of 1.82 Å for 367  $C^\alpha$  atoms. This difference is due to the opening of the substrate-binding cleft in the *SpnHyal*-PEGMME structure and was analyzed further using Dyndom<sup>22</sup> to study dynamic rather than structural domains. Dyndom defines two dynamic domains, the first containing just residues from the  $\alpha$ -domain, 173–378 and 420–431, the second composed of the entire  $\beta$ -domain as well as portions of the  $\alpha$ -domain, residues 379–419 and 432–888. In the *SpnHyal*-PEGMME structure these dynamic domains are differently positioned with respect to each other so that the substrate binding cleft is significantly wider than in all earlier pneumococcal Hyal structures, including the complexes with substrates and/or products. The extent of rotation of the two domains around hinge regions is measured as 15.5° by Dyndom. At the catalytic site, the  $C^\alpha$ – $C^\alpha$  distance between two catalytic residues Asn349 and Tyr408 (**Figure 2(b)**) is ~7.7 Å in the original Hyal structures but is 9.7 Å in the *SpnHyal*-PEGMME structure. Further from the catalytic center, the cleft opening is much more dramatic, the  $C^\alpha$ – $C^\alpha$  distance between Asn290, a neighbor of hydrophobic patch residues (**Figure 2(b)**), and its nearest neighbor across the cleft Asn480 (not shown), is 8.4 Å in *SpnHyal*, but 14.9 Å in *SpnHyal*-PEGMME. This widening of the cleft is reminiscent of that observed earlier in the *SagHyal* structure.<sup>7</sup> Indeed the dynamic domains defined for the *SpnHyal* versus *SagHyal* comparison are very similar to those described for the comparison of *SpnHyal* with *SpnHyal*-PEGMME. However, the *SpnHyal*-PEGMME cleft opening differs from that observed in the *S. agalactiae* enzyme in two important respects. First, it is still wider, 15.5° compared to 13.1°. The second difference between the *SpnHyal*-PEGMME and *SagHyal* structures is that the cleft opening of the former is accompanied by another smaller difference, the rotation of the side of the cleft close to the  $\beta$ -domain in the HA reducing direction (**Figures 2(b) and 3(a)**). Such a rotation was not observed when comparing the original *SpnHyal* and *SagHyal* structures. It, too, is evidence in support of enzyme flexibility that facilitates catalysis (see below).

In the *SpnHyal*-MALONATE structure the structural differences are subtler than in that of *SpnHyal*-PEGMME. The whole *SpnHyal*-MALONATE structure may be superimposed on the *SpnHyal* structure with an rmsd of 0.6 Å for 719  $C^\alpha$  residues. Structural differences are greater in the  $\alpha$ -domain (rms of 0.7 Å for 367 superimposed  $C^\alpha$  atoms) than in the  $\beta$ -domain (rms of 0.5 Å for 352 superimposed  $C^\alpha$  atoms). The average cleft width is more in line with that of *SpnHyal*, but is somewhat wider. Using the same benchmark as above, the inter- $C^\alpha$  distance between Asn290 and Asn480 is 9.6 Å (compared to ~8.4 Å in earlier structures).

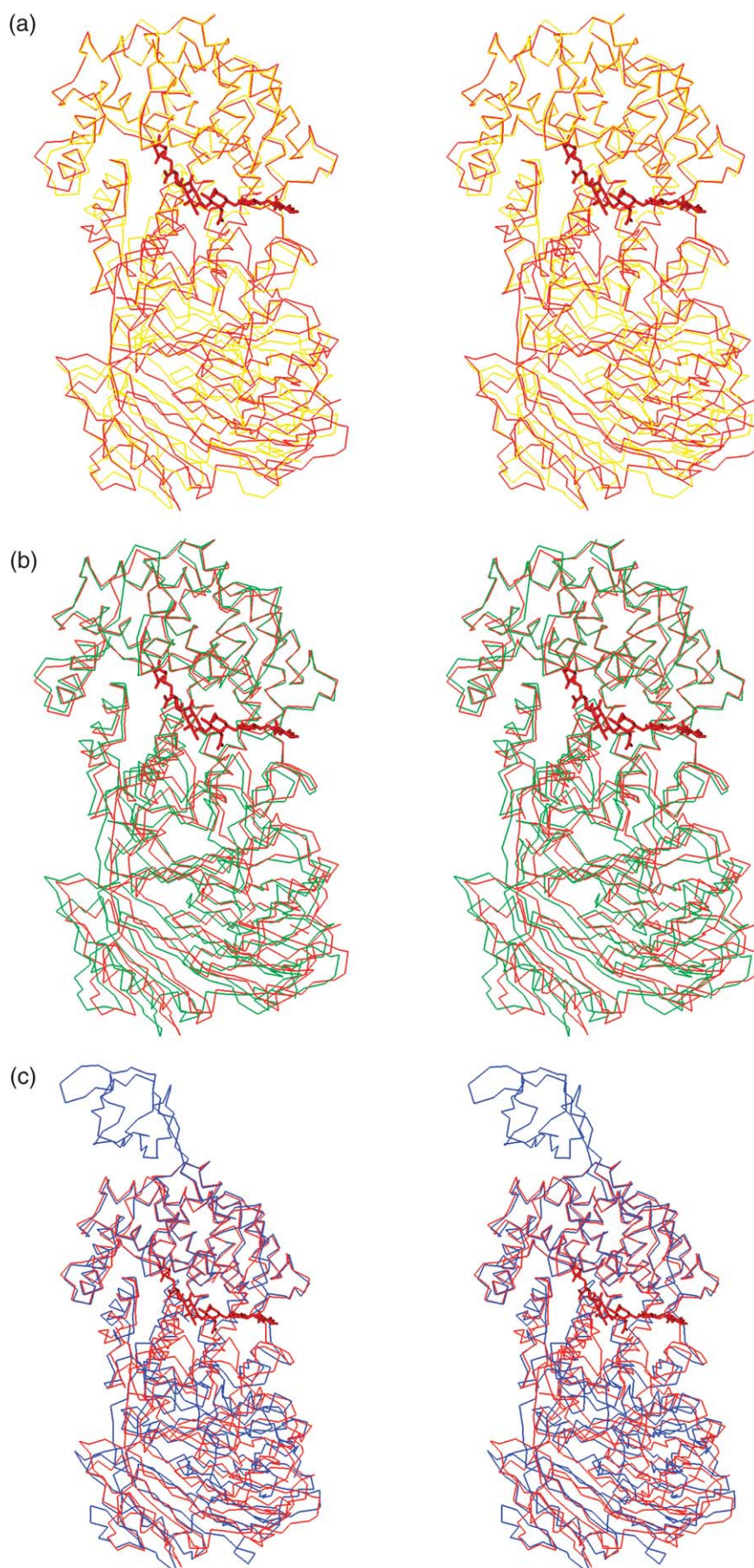
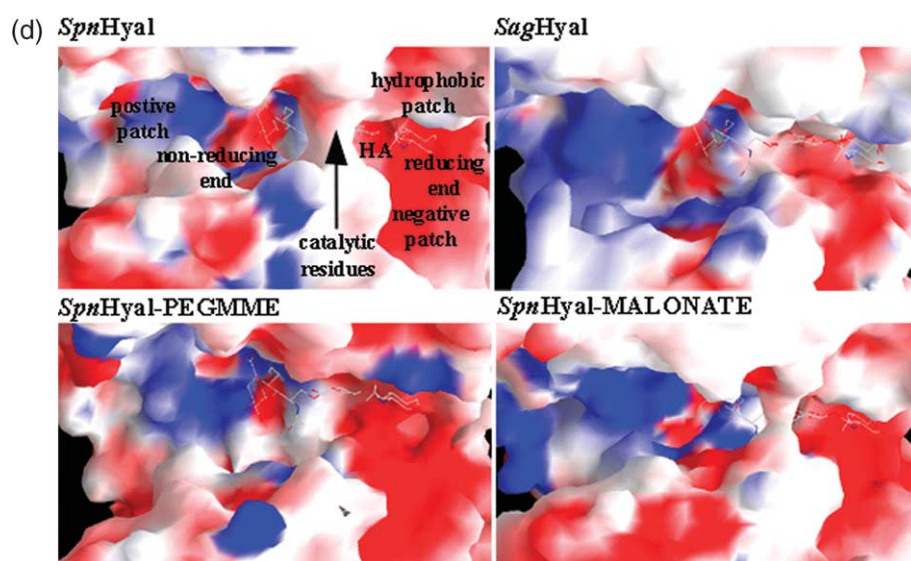


Figure 3 (legend next page)



**Figure 3.** Structural alignment and comparison of the *S. pneumoniae* Hyal-PEGMME and Hyal-MALONATE structures of the enzyme in new conformations with structures of *S. pneumoniae* Hyal (*holo*) and *S. agalactiae* enzyme. The structures were aligned with *SpnHyal* using the most similar portions of the  $\alpha$ -domain for the structures involved (color-coding for structures: *SpnHyal-holo* (red) with bound HA hexasaccharide substrate (ball and sticks, red), *SpnHyal-PEGMME* (yellow), *SpnHyal-MALONATE* (green). The orientations of structures in (a)–(d) are similar to the orientation of the active residues in Figure 2(b). (a) Stereo view of *SpnHyal-PEGMME* aligned with *SpnHyal-holo*. (b) Stereo view of *SpnHyal-MALONATE* aligned with *SpnHyal-holo*. (c) Stereo view of *SpnHyal-holo* aligned with *SagHyal*. The *S. agalactiae* enzyme has the  $\beta$ -sheet spacer domain at its N-terminal end and a wider HA-binding cleft. (d) Side-by-side view of potential distribution inside the catalytic cleft for representative structures. The location of catalytic, hydrophobic, negative, and positive patches as well as the reducing/non-reducing ends of the HA substrate are labeled. The electrostatic potential magnitude is held constant for all enzyme panels, and is color-coded as follows: blue, positive; red, negative.

Also a small rotation of the cleft side that is in the proximity of the  $\beta$ -domain is evident. For this structure, however, the rotation is towards the non-reducing direction of the bound HA substrate in the opposite direction than that of the *SpnHyal-PEGMME* conformation (Figures 2(b) and 3(b)). There are relatively large differences in and around the active site particularly in the vicinity of a bound malonate ion (from the crystallization solution). The differences related to the bound malonate ion are presumed to be non-physiological and may have been induced by crystallization conditions. They do not contribute to the rotation described above.

For further analysis, the new structures were superimposed (using the  $\alpha$ -domain) on the *SpnHyal-holo* enzyme structure, presumably the conformation most closely resembling the catalytically competent state. The relative positions of specific regions within the catalytic site, the hydrophobic patch, the positive patch and the negative patch<sup>8</sup> (Figure 2(b)), were analyzed. It appears that the only part of the enzyme that does not vary significantly is the hydrophobic patch. The positions of Trp291, Trp292, and Phe343 (and equivalents for *SagHyal*) in all structures of bacterial Hyals superimpose to within 1 Å. In sharp contrast, in the *SpnHyal-PEGMME* structure relative to the *SpnHyal-holo* enzyme structure, Tyr408 and His399, are shifted by  $\sim 4.4$  Å along the increased width of the cleft, whereas Asn349 is translocated by  $\sim 4.5$  Å in the reducing end

direction of the substrate and along the cleft axis (Figure 2(b)). The same distances for the *SpnHyal-MALONATE* structure comparison are around  $\sim 1.0$  Å, showing its resemblance to the *SpnHyal-holo* structure in this respect. Interestingly, for *SagHyal* distances of  $\sim 4.4$  Å for Tyr408 and His399 are measured corresponding to a cleft opening, but Asn349 positions are similar to within 1 Å. Finally, the residues of negative patch Glu388, Asp398, and Thr400 appear to move the most. Compared to *SpnHyal-holo*, in the *SpnHyal-PEGMME* structure this patch is essentially translocated as a whole by  $\sim 4.8$  Å in the reducing HA direction along the cleft. The other two structures show opposite translocation towards the non-reducing HA end by  $\sim 1.8$  Å for *SpnHyal-MALONATE* and  $\sim 3.5$  Å for *SagHyal*.

Analysis of shape and distribution of electrostatic potential within the cleft highlights marked differences. The electrostatic potential does not vary significantly during enzyme flexing leaving unchanged the positions of the positive patch (at the non-reducing end) and the negative patch sited at the reducing end (Figure 3(d)).<sup>13,17</sup> Presumably their respective functions of continuous substrate binding (positive patch) and the release of the disaccharide product (negative patch) remain important throughout the catalytic cycle. In contrast, the shape of the cleft varies significantly among all structures. In the closed conformation of *SpnHyal-holo* it is even covered by a bridge over

the catalytic residues forcing the substrate to thread through the cleft for catalysis. In *SpnHyal*-PEGMME the cleft is widely open, in a conformation that would facilitate processivity by allowing easy substrate translocation towards the reducing end.<sup>8,18</sup> The remaining structures appear to illustrate intermediate cleft opening/closing stages, in the *SpnHyal*-MALONATE structure the cleft is nearly closed and in the *SagHyal* structure it is nearly open.

### Bacterial Hyals are flexible

The Hyal enzyme exhibits a high flexibility that facilitates the mechanism of action of this enzyme. A previous flexibility analysis revealed three major types of motion of the enzyme (Figure 4):<sup>13,17</sup> (i) a rotation/twisting motion of the whole  $\alpha$ -domain relative to the top half of the  $\beta$ -domain (Figure 4(a)); (ii) an opening/closing domain motion influencing the width of substrate-binding cleft (Figure 4(b)); and (iii) the mobility of loops from the  $\beta$ -domain, relative to the  $\alpha$ -domain, that are involved in the formation of the side of the cleft resulting in opening/closing of the access/entrance to the cleft (Figure 4(c)). For the sake of clarity we maintain the same numbering of motions in the present work.

Here, we present multi-nanosecond molecular dynamics simulations of *SpnHyal-holo* and *SpnHyal-apo* in explicit solvent. These represent a significant advance on the earlier computational analyses made with the computationally efficient but quantitatively less rigorous CONCOORD program.<sup>17,23</sup> These simulations confirm the results from the previous flexibility analysis and reveal these three types of domain motion as the principal modes of collective fluctuation as derived from a principal components analysis (PCA). In both molecular dynamics (MD) simulations the opening/closing mode of the cleft (ii) is the largest-amplitude fluctuation. As can be seen in Figure 5, the extent of the opening is in line with the X-ray structures, with the most open structures being even more widely opened than the *SagHyal* and *SpnHyal*-PEGMME crystal structures. The opening motion is found to be reversible in the simulations, with the most closed conformations demonstrating the same degree of closure as the X-ray structures of the closed enzyme, *SpnHyal*. Along this opening mode, the *SagHyal* structure, and others from the same species (green), exhibit a slightly higher degree of opening than the *SpnHyal*-PEGMME structure.

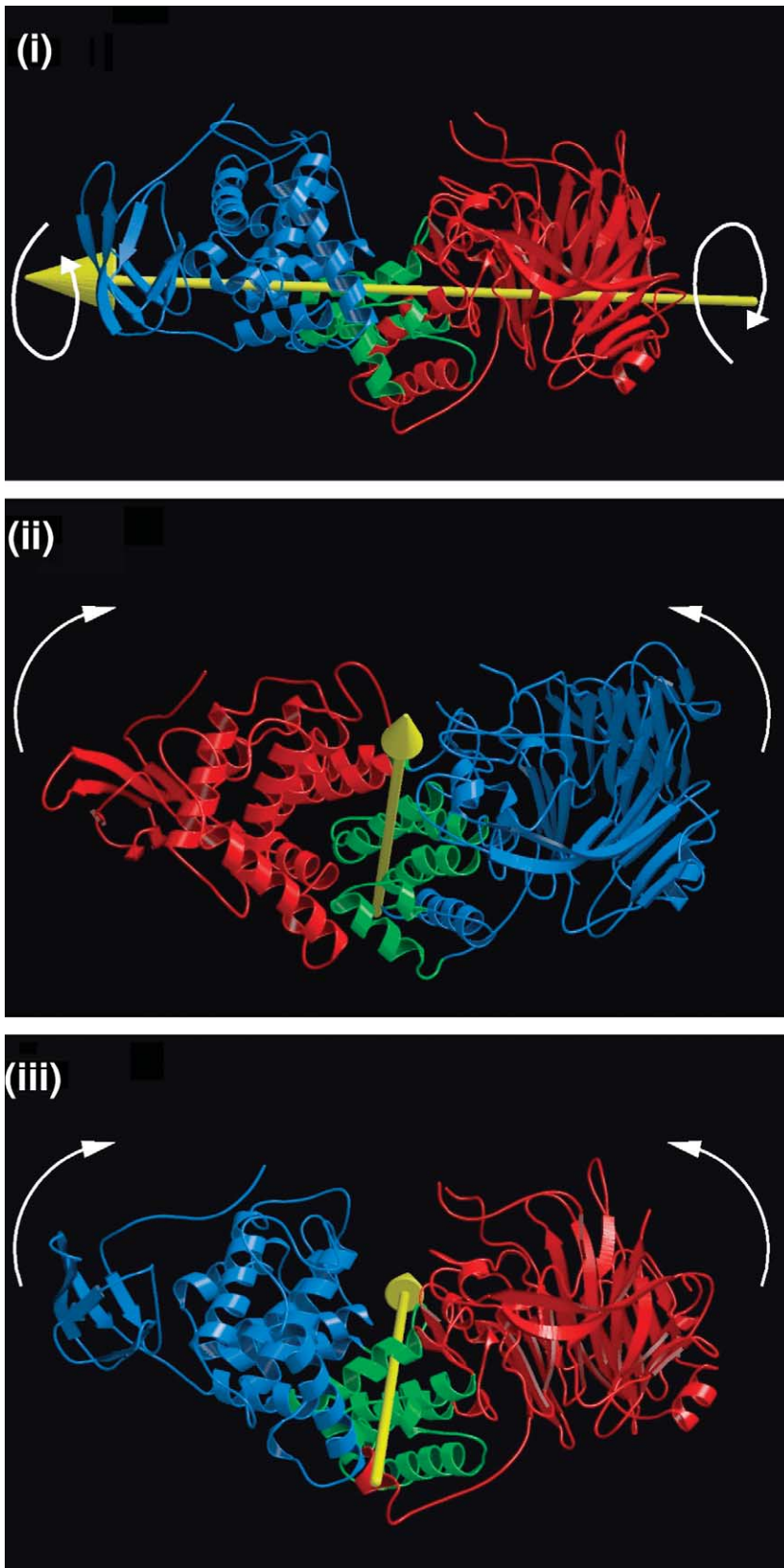
An rmsd analysis was carried out to quantify the structural differences. The rmsd between the crystallographic starting structure of the simulations (PDB code 1loh (Table 1), the "closed" conformation) and the *SpnHyal*-PEGMME structure (the "open" conformation) is 2.2 Å. During the simulation of *SpnHyal-apo* the rmsd, compared to the PEGMME structure, decreases to as little as 1.2 Å. Similarly for the simulation of *SpnHyal-holo* it is reduced to as little as 1.2 Å. This confirms that in

both simulations the opening mode is significantly sampled. Interestingly, the MD simulation including a hyaluronan hexasaccharide does not sample the more open configurations to the same extent as the simulation of *SpnHyal-apo* within the same time span of 50 ns (Figure 5(a)), suggesting that the substrate locks the protein in a more closed conformation. The maximal amplitude of the opening of the cleft of *SpnHyal-apo* is observed to be up to 12.8 Å (measured as the difference between the C $^{\alpha}$ -C $^{\alpha}$  separations of Asn231 and Gly769 in the most open and least open trajectory snapshots) In the simulation of *SpnHyal-holo* the extent of the opening of the cleft, measured in the same way, is observed only up to 6.5 Å (see Figure 5(a) and Table 3). This is further quantified by the mean projection along this mode for the *SpnHyal-apo* simulation of  $0.64 \pm 1.16$  (standard deviation) while for the *SpnHyal-holo* simulation it is  $0.01 \pm 0.51$ . For comparison, the Asn231-Gly769 distances of the *SpnHyal-holo*, *SpnHyal*-PEGMME and *SagHyal* structures are 14.0 Å, 22.1 Å and 21.6 Å.

The second-largest amplitude domain motion observed in the simulations corresponds to the twisting mode (i) from the original flexibility analysis. This mode results in a reciprocal motion of the two domains with respect to the cleft, and might therefore enable an effective translocation of the substrate along the cleft. This presumption is supported by the extent of the twisting motion in the simulations. In the simulation of *SpnHyal-apo* the  $\alpha$ -domain is observed to be shifted with respect to the  $\beta$ -domain due to this twisting mode by up to 10.92 Å. This extent was measured as the separation between the C $^{\alpha}$  atoms of Asp340 for the superimposed most twisted and least twisted snapshots from the trajectory. Strikingly, this is also approximately the length of a disaccharide unit of the hyaluronan substrate. Also along this mode, the simulated motions agree closely with the crystallographic structural differences.

Along this twisting mode (i), the *SpnHyal*-PEGMME structure represents a structural difference in the opposite direction as the *S. agalactiae* X-ray structures, *SagHyal*. The MD simulations span the complete range sampled by the X-ray structures, suggesting that this mode also reflects an intrinsic mode of flexibility of the protein. These results also indicate that, although both the opening/closing (ii) and the twisting mode (i) are accessible to *S. pneumoniae* Hyal, these two modes appear to be coupled during the simulation, leaving the *S. agalactiae* X-ray structures (green) unsampled during the simulations. Thus, although the *S. pneumoniae* and the *S. agalactiae* X-ray structures are both reflecting the inherent flexibility of this family of enzymes, part of the structural difference presumably results from the apparent sequence differences.

Mode (iii), corresponding to a bending of the two domains that influences the width of the cleft entrance/exit, is also the third-largest amplitude mode from the MD simulations. Along this mode, the *SpnHyal* structure samples an intermediate



**Figure 4.** The three principal modes of collective fluctuation for bacterial Hyals. The motions described are marked by arrows. The yellow arrow corresponds to a rotation axis for the domain motion described and the white arrows are purely descriptive. The different colors highlight independently flexible portions of the enzyme (red and blue domains and their green linker; these are different than the structural  $\alpha$  and  $\beta$ - domains and their linker. Motion (i) describes a twisting motion of the two domains, putatively enabling substrate translocation across the cleft; (ii) corresponds to a closure mode, opening and closing the substrate cleft; (iii) describes a bending mode putatively facilitating substrate access and product release at the gorge entry and exit, respectively.

position and the *SpmHyal*-PEGMME structure and *SagHyal* structures represent a conformation with the entrance opened and the exit closed. The MD simulations demonstrate that this motion is

sampled to the same extent in the opposite direction, thereby enabling opening the exit and closing the entrance to the cleft. The amplitude of this motion was measured as the difference between





**Table 3.** The extent of the dynamic domain motions in *apo* and *holo*-HL simulations

Eigenvector/motion description	<i>SpnHyal-apo</i> (Å)	<i>SpnHyal-holo</i> (Å)
Eigenvector (i) (opening and closing of the cleft)	12.8	6.5
Eigenvector (ii) (twisting of the $\alpha$ -domain)	10.9	5.9
Eigenvector (iii) (opening and closing of the access to the cleft)	9.1	5.5

See the text for the details of the measurements.

the C $^{\alpha}$ -C $^{\alpha}$  separations of Asn341 and Asn580 in the trajectory snapshots with most open and least open cleft entrances in the same way as for the mode (ii). This distance was found to be maximally 9.08 Å in the simulation of *SpnHyal-apo*, whereas for *SpnHyal-holo* the amplitude of this cleft-access mode was found to be maximally 5.47 Å. Thus, this motion is significantly better explored in the *SpnHyal-apo* structure.

## Discussion

We present here novel structures of *SpnHyal*. They were obtained using different crystallization conditions than those hitherto employed. As such, it could be argued that they do not necessarily reflect physiologically relevant conformations. However, several factors lead us to believe in the validity and relevance of the structures. First, although precipitant can certainly influence the crystallized conformation, it is unlikely that it could “force” the molecule into an energetically unfavorable conformation. In fact, conformational diversity in structures of the same protein has often been correlated with functionally important motions. Thus, the conformations captured in the *SpnHyal*-PEGMME and *SpnHyal*-MALONATE structures can reasonably be supposed to be accessible to the enzyme and hence of potential importance to its *in vivo* function. In this regard, it should be noted that the pH values at which the structures crystallized, pH 6.5 and 6.6 for the *SpnHyal*-PEGMME and *SpnHyal*-MALONATE structures, respectively, are neither far from the enzyme’s pH optimum of  $\sim$ pH 6.0–6.5, nor very different from the pH 6.0 at which the enzyme has previously been crystallized. The pneumococcal Hyal is a very resilient enzyme and remains active under variety of conditions including reasonable amounts of ammonium sulfate (original crystallization medium for this enzyme) or sodium malonate salts, or PEG, including PEGMME (not shown). As expected, however, at high concentrations of these crystallizing molecules the enzyme activity is compromised. Regardless, the Hyal molecules within crystalline lattice or once such crystals are re-dissolved under standard assay conditions<sup>9,10</sup> they contain molecules capable of full activity.<sup>13,17</sup> Importantly, it is commonly the

case that the molecular dynamics simulations can reproduce modes of structural variability observed in sets of crystal structures and, indeed, we show that to be the case for Hyal. Finally, the most radically different conformation of *SpnHyal* described here, *SpnHyal*-PEGMME, bears a clear resemblance to a previous structure of *SagHyal* making less likely the idea that it is an artifact of no physiological relevance.

With the reliability of the new structures established, their implications for our understanding of Hyal flexibility and function will be discussed. Based on computer simulations studies, a cleft opening/closing motion (ii) has been predicted to be essential for catalysis. The new *SpnHyal*-PEGMME structure now provides the first experimental evidence that directly points to the ability of Hyal to achieve this earlier predicted open conformation. Such an open conformation was previously observed for *SagHyal* but the different species origin of that protein left a doubt as to the universality of the flexibility leading to the conformation. The new structure reveals that this cleft-widening motion is an intrinsic mechanistic property of the enzyme, independent of the protein source.

The structural differences between all Hyals may be correlated to their modes of action, as revealed by kinetic analysis with a variety of substrates.<sup>14,24,25</sup> As mentioned, biochemical data demonstrate that Hyal cleavage of HA, its primary substrate, proceeds *via* an initial endolytic cleavage until the average HA chain decreases to approximately 300 kDa, at this stage another mode of action based on processive, exolytic HA degradation of one disaccharide at a time initiates. As HA chains reach  $\sim$ 100 disaccharides in length the rapid exolytic and processive activity producing a solely unsaturated disaccharide unit as the end degradation product is the primary mode of action.<sup>13,19,20,26</sup> Such behavior is directly related to HA aggregation when it is a large polymer, but such aggregation decreases as the HA size diminishes and below 40 kDa (100 disaccharides)<sup>27,28</sup> HA is thought to be not aggregated, resulting in processive action of Hyal. Bacterial Hyals degrade chondroitin substrates by a purely endolytic process.

With the benefit of the new crystal structures present here, and the wholly consistent multi-nanosecond MD simulations reported, we feel able to propose a consistent underlying molecular mechanism of processive degradation of substrate(s) by bacterial Hyals. The process would consist of the following steps. Once the access to the cleft is open (motion (iii)) and the cleft is widely open (motion (ii)) the substrate is free to bind in the cleft. Upon binding, the entrance closes (motion (iii)) and the cleft width diminishes together with catalytic residues moving into their functional positions (motion (ii)). Once they perform catalysis, the residues responsible for product release move in to repel this product (motion (ii)). For further

processive exolytic degradation of one substrate disaccharide at a time, the gate remains closed (motion (iii)) preventing the substrate escape from the cleft and the cleft widens (motion (ii)). The substrate-bound enzyme then undergoes the twisting motion (i), by which the substrate is translocated in the direction of its reducing end by  $\sim 10$  Å, also the length of a disaccharide. This would be followed by motion (ii), cleft closure and catalysis then cleft opening and product release, and repetition of the cycle until the substrate is fully degraded in the processive manner. In the endolytic degradation mode, both cleaved products would be released from the cleft after the catalysis, a process facilitated by gate opening (motion (iii)) and widening of the cleft (motion (ii)). The type of degradation is substrate specific and depends on the length and aggregation of the substrate.<sup>18</sup> Further experiments complemented by additional simulations will be required to study the individual steps in detail to verify the proposed mechanism.

## Conclusions

We have studied the structural basis of the molecular mechanism underlying the cleavage of substrates by bacterial Hyals by X-ray crystallography, as well as by molecular dynamics simulations. It is not common for enzymes to crystallize under different conditions yielding different structures of the same enzyme. In the case of Hyals from two streptococcal species, *S. pneumoniae* and *S. agalactiae*, crystals in a number of structural conformations were obtained. These conformations were found to closely agree with the dynamic behavior of the enzyme as studied by molecular dynamics simulations and were convincingly associated with the enzyme's mechanism of processive hyaluronan degradation at the molecular level. Obtaining the new structures allowed for the validation of the computational models of enzyme action.

In addition, a basic property of the nature of the processive mechanism of degradation was revealed by diverse structural approaches. The flexibility of Hyal appears to directly facilitate processivity of this enzyme by threading the linear polymeric substrate(s) through the enzymatic/catalytic cleft. The significant insights into detailed mechanisms on a microscopic level is not common, and for Hyals they yielded a deeper understanding of this process. There are numbers of aspects of this enzyme's functional behavior which are still not understood, such as what is the significance of the generation of disaccharide products, why nature evolved to produce an unsaturated product only by the action of bacterial and not animal Hyals, why the disaccharides are presumably used as a source of carbon and energy by bacteria, and what is the mechanism of this product presumed transport into bacterial cytoplasm? Further studies should

facilitate elucidation of these and other properties of Hyals and their functionality.

## Materials and Methods

### Crystallization, and diffraction data collection

The wild-type *S. pneumoniae* Hyal, in its fully functional truncated form composed of the catalytic and C-terminal domains only, was produced as reported.<sup>9,10</sup> The protein was concentrated to 5 mg/ml in 10 mM Tris-HCl buffer (pH 7.4), 150 mM NaCl, 2 mM EDTA. For crystallization the hanging drop vapor diffusion method was employed<sup>29</sup> and equal volumes (1  $\mu$ l each) of the enzyme and reservoir solution were utilized as reported.<sup>10</sup> Two new crystal forms were obtained. These crystals grew within several months and were of different morphology than of those of the wild-type enzyme reported.<sup>10</sup> The first set of crystals was obtained employing 30% (w/v) polyethylene glycol monomethyl ether (PEGMME) 2000, 0.1 M Mes buffer (pH 6.5), and 0.2 M ammonium sulfate. The second set of crystals was obtained utilizing 70% (w/v) saturated malonic acid (pH 7.6) and 0.1 M sodium cacodylate buffer (pH 6.6).

The crystals were cryo-protected and frozen using standard methodology. The diffraction data were collected using a synchrotron source, beamline 5.0.1 of the Berkeley Center for Structural Biology, Advanced Light Source, Lawrence Berkeley National Laboratory. An ADSC Quantum 4u CCD detector was used and the data were collected in oscillation mode.<sup>30</sup> The new crystals were not isomorphous to those of the original crystals of this wild-type enzyme.<sup>10</sup> The PEGMME crystals yielded X-ray diffraction data at 2.8 Å, and indexed to the *P21* space group with unit cell dimensions  $a=86.4$  Å,  $b=83.3$  Å,  $c=98.7$  Å, and  $\gamma=98.8^\circ$ . The malonate crystals yielded a 3.3 Å diffraction data set, space group *C2*, and unit cell dimensions of  $a=117.7$  Å,  $b=101.1$  Å,  $c=85.2$  Å, and  $\gamma=125.1^\circ$ . The reasons for the relatively poor diffraction by the new crystal form are unclear. Heterogeneity in the protein sample seems unlikely, since Hyal readily and reproducibly crystallizes in the original crystal form. Neither new crystal form is characterized by a high solvent content, 43% for the malonate-derived crystals and 50% for the PEGMME crystals. The crystal took a very long time to grow, and conceivably the structural flexibility of the enzyme is reflected in the crystal quality and the resulting X-ray diffraction data.

### Structure solution and refinement

The structures were solved by molecular replacement methods employing the MOLREP program<sup>31</sup> and utilizing the 1.56 Å crystal structure of the wild-type *S. pneumoniae* Hyal structure (*SpnHyal*) as a search model. This structure was then divided into smaller portions based on the structural and dynamic domains previously identified.<sup>13,17</sup> This approach was necessary to achieve a successful solution of the *SpnHyal*-PEGMME structure. Solutions were further refined using rigid body refinement in CNS.<sup>32</sup> Refinement was principally done with CNS<sup>32</sup> but REFMAC5,<sup>33</sup> in combination with ARP/WARP,<sup>34</sup> proved to be useful at the first stage of positional refinement of *SpnHyal*-MALONATE. Computational refinement was alternated with manual rebuilding with O.<sup>35</sup> All data were used throughout with no intensity or

sigma-based cut-offs applied. SigmaA-weighted map coefficients<sup>36</sup> were used throughout. Maximum likelihood functions were used throughout. An  $R_{\text{free}}$  value,<sup>37</sup> calculated from 5% of reflections set aside at the outset, was used to monitor the progress of refinement. The asymmetric unit of the *SpnHyal*-PEGMME crystal form contains two Hyal chains. During initial refinement, these were constrained to identical conformations with CNS. Later, the constraints were replaced by restraints and, where density suggested conformational differences between the chains, specific regions were exempted entirely from the restraints. During the final stages of refinement exempted residues were 385–388, 420–433, 435, 444, 492–493, 518, 524 and 527–540. At a late stage, water molecules were placed into  $3\sigma$  positive peaks in  $|F_o - F_c|$  maps when density was also evident in  $|2F_o - F_c|$  maps and suitable hydrogen bonding partners were available. Sulfate and malonate ions derived from the crystallization solutions were modeled into suitably shaped regions of electron density. Temperature factors were refined first on a per-residue basis, and then per-atom for a limited number of iterations guided by the behavior of the  $R_{\text{free}}$  value. Final statistics for the two structures are presented in Table 2. The water molecules in the *SpnHyal*-PEGMME structures refined to atypically low values. The reasons for this are unclear but may include (1) the fact that they were picked conservatively and so the final group represents the best-ordered subset and (2) some of the water molecules are in fact heavier cations or anions, hard to distinguish at the available resolution. Programs of the CCP4 package<sup>38</sup> were used for all data and structure manipulations and structural superpositions were performed with LSQMAN.<sup>39</sup>

Structural alignments and superpositions for comparison purposes were performed based on the alignment of the catalytic  $\alpha$ -domain only. Structural placement and positioning of the HA substrate within the cleft was performed based on the superposition of individual structures with that of the published *S. pneumoniae* *holo*Hyal-HA hexasaccharide complex structure.<sup>8</sup> The comparison with the *S. agalactiae* Hyal-HA complex was based on the alignment (based on C $^{\alpha}$  atoms of the  $\alpha$ -domain only) with the published structure of its complex with HA hexasaccharide.<sup>17</sup> All structural Figures were produced with PyMol,<sup>†</sup> Ribbons,<sup>40</sup> or Grasp.<sup>41</sup>

### Molecular dynamics simulations

MD simulations of 50 ns each were carried out on *SpnHyal-*apo** and *SpnHyal-*holo** (see Table 1). The simulations were carried out using the GROMACS simulation suite,<sup>42</sup> employing the optimized potentials for liquid simulations, all atom (OPLS-AA) force field.<sup>43</sup> Hyaluronan parameters were derived by modifying the molecule-specific dihedral parameters in the OPLS-AA force field for the carbohydrate substrate.<sup>44</sup> The enzyme was solvated in a periodic box of TIP4P water molecules, resulting in a simulation size of 77,278 and 83,542 atoms for the *SpnHyal-*apo** and *SpnHyal-*holo** simulations, respectively. LINCS<sup>45</sup> was used to constrain covalent bonds, allowing a time step of 2 fs. The temperature was kept constant at 300 K by separately coupling (time step  $\tau=0.1$  ps) protein and solvent to an external temperature bath.<sup>46</sup> Likewise, the pressure was kept at 1 bar ( $\tau=1$  ps)

by using a Berendsen barostat.<sup>46</sup> Electrostatic interactions were calculated using the smooth particle mesh Ewald (PME) method.<sup>47</sup>

### Other methods

The Dyndom<sup>22</sup> method was used to analyze domain motions. The enzyme concentration was determined by UV absorption at 280 nm using the molar extinction coefficient calculated based on the native or mutant *S. pneumoniae* Hyal amino acid residue sequence data.<sup>9,10</sup> The calculated molar extinction coefficient was 127,090 for the wild-type truncated enzyme used.

Principal component analysis (PCA) is a tool for filtering large-scale collective fluctuations from MD trajectories or a set of experimental structures. The method is based on the diagonalization of the covariance matrix of atomic fluctuations, which yields a set of eigenvalues and eigenvectors. The eigenvectors indicate directions in a  $3n$ -dimensional space (with  $n$  being the number of atoms of the protein) and describe concerted fluctuations of the atoms. The eigenvalues reflect the magnitude of the fluctuation along of the respective eigenvectors. To compare the collective motions of two simulations of *SpnHyal-*apo** and *SpnHyal-*holo**, the PCA was first performed on crystallographic data available. The available dataset includes X-ray structures of Hyal from both *S. pneumoniae* and *S. agalactiae* in both *apo* and *holo* forms and comprises the structures in Table 1 along with structures having PDB codes 1c82, 1i8q, 1lxx and 1lxm (other substrate complexes) and 1f9g (a vitamin C complex). The two simulations were projected onto this set of eigenvectors for comparison.

### Protein Data Bank accession codes

The coordinates and structural factors of both structures, *SpnHyal*-PEGMME and *SpnHyal*-MALONATE, have been deposited in the Protein Data Bank with accession numbers 2brw, 2brv, and r2brwsf, r2brvsv, respectively.

### Acknowledgements

The authors thank Dr R. Keith Henderson for his help and assistance with the diffraction data collection. The diffraction data were collected at beamline 5.0.1 of the Berkeley Center for Structural Biology, Advanced Light Source, Lawrence Berkeley National Laboratory.

### References

1. Linker, A., Meyer, K. & Hoffman, P. (1956). The production of unsaturated uronides by bacterial hyaluronidases. *J. Biol. Chem.* **219**, 13–25.
2. Berry, A. M., Lock, R. A., Thomas, S. M., Rajan, D. P., Hansman, D. & Paton, J. C. (1994). Cloning and nucleotide sequence of the *Streptococcus pneumoniae* hyaluronidase gene and purification of the enzyme from recombinant *Escherichia coli*. *Infect. Immun.* **62**, 1101–1108.

† <http://pymol.sourceforge.net>

3. Pallen, M. J., Lam, A. C., Antonio, M. & Dunbar, K. (2001). An embarrassment of sortases—a richness of substrates? *Trends Microbiol.* **9**, 97–102.
4. Ton-That, H., Mazmanian, S. K., Faull, K. F. & Schneewind, O. (2000). Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Sortase catalyzed *in vitro* transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates. *J. Biol. Chem.* **275**, 9876–9881.
5. Schneewind, O., Fowler, A. & Faull, K. F. (1995). Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. *Science*, **268**, 103–106.
6. Rigden, D. J. & Jedrzejewski, M. J. (2003). Genome-based identification of a carbohydrate binding module in *Streptococcus pneumoniae* hyaluronate lyase. *Proteins: Struct. Funct. Genet.* **52**, 203–211.
7. Li, S. & Jedrzejewski, M. J. (2001). Hyaluronan binding and degradation by *Streptococcus agalactiae* hyaluronate lyase. *J. Biol. Chem.* **276**, 41407–41416.
8. Li, S., Kelly, S. J., Lamani, E., Ferraroni, M. & Jedrzejewski, M. J. (2000). Structural basis of hyaluronan degradation by *Streptococcus pneumoniae* hyaluronate lyase. *EMBO J.* **19**, 1228–1240.
9. Jedrzejewski, M. J., Mewbourne, R. B., Chantalat, L. & McPherson, D. T. (1998). Expression and purification of *Streptococcus pneumoniae* hyaluronate lyase from *Escherichia coli*. *Protein Expr. Purif.* **13**, 83–89.
10. Jedrzejewski, M. J., Chantalat, L. & Mewbourne, R. B. (1998). Crystallization and preliminary X-ray analysis of *Streptococcus pneumoniae* hyaluronate lyase. *J. Struct. Biol.* **121**, 73–75.
11. Ponnuraj, K. & Jedrzejewski, M. J. (2000). Mechanism of hyaluronan binding and degradation: structure of *Streptococcus pneumoniae* hyaluronate lyase in complex with hyaluronic acid disaccharide at 1.7 Å resolution. *J. Mol. Biol.* **299**, 885–895.
12. Nukui, M., Taylor, K. B., McPherson, D. T., Shigenaga, M. K. & Jedrzejewski, M. J. (2003). The function of hydrophobic residues in the catalytic cleft of *Streptococcus pneumoniae* hyaluronate lyase. Kinetic characterization of mutant enzyme forms. *J. Biol. Chem.* **278**, 3079–3088.
13. Jedrzejewski, M. J., Mello, L. V., de Groot, B. L. & Li, S. (2002). Mechanism of hyaluronan degradation by *Streptococcus pneumoniae* hyaluronate lyase. Structures of complexes with the substrate. *J. Biol. Chem.* **277**, 28287–28297.
14. Kelly, S. J., Taylor, K. B., Li, S. & Jedrzejewski, M. J. (2001). Kinetic properties of *Streptococcus pneumoniae* hyaluronate lyase. *Glycobiology*, **11**, 297–304.
15. Rigden, D. J., Galperin, M. Y. & Jedrzejewski, M. J. (2003). Analysis of structure and function of putative surface-exposed proteins encoded in the *Streptococcus pneumoniae* genome: a bioinformatics-based approach to vaccine and drug design. *Crit. Rev. Biochem. Mol. Biol.* **38**, 143–168.
16. Jedrzejewski, M. J. & Chantalat, L. (2000). Structural studies of *Streptococcus agalactiae* hyaluronate lyase. *Acta Crystallog. sect. D*, **56**, 460403.
17. Mello, L. V., de Groot, B. L., Li, S. & Jedrzejewski, M. J. (2002). Structure and flexibility of *Streptococcus agalactiae* hyaluronate lyase complex with its substrate. Insights into the mechanism of processive degradation of hyaluronan. *J. Biol. Chem.* **277**, 36678–36688.
18. Rigden, D. J. & Jedrzejewski, M. J. (2003). Structures of *Streptococcus pneumoniae* hyaluronate lyase in complex with chondroitin and chondroitin sulfate disaccharides: insights into specificity and mechanism of action. *J. Biol. Chem.* **278**, 50596–50606.
19. Jedrzejewski, M. J. (2002). Three-dimensional structures of hyaluronate lyases from *Streptococcus* species and their mechanism of hyaluronan degradation. In *Science of Hyaluronan Today* (Hascall, V. C. & Yanagishita, M., eds), Tokyo, Japan.
20. Jedrzejewski, M. J. (2002). Mechanisms of polysaccharide degradation by bacterial enzymes: degradation of hyaluronan. In *Recent Research Developments in Biophysics and Biochemistry* (Pandali, S. G., ed.), vol. 2, pp. 197–225, Research Signpost, Trivandrum, India.
21. Li, S., Taylor, K. B., Kelly, S. J. & Jedrzejewski, M. J. (2001). Vitamin C inhibits the enzymatic activity of *Streptococcus pneumoniae* hyaluronate lyase. *J. Biol. Chem.* **276**, 15125–15130.
22. Hayward, S. & Berendsen, H. J. (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins: Struct. Funct. Genet.* **30**, 144–154.
23. de Groot, B. L., Vriend, G. & Berendsen, H. J. (1999). Conformational changes in the chaperonin GroEL: new insights into the allosteric mechanism. *J. Mol. Biol.* **286**, 1241–1249.
24. Jandik, K. A., Gu, K. & Linhardt, R. J. (1994). Action pattern of polysaccharide lyases on glycosaminoglycans. *Glycobiology*, **4**, 289–296.
25. Gu, K., Linhardt, R. J., Laliberte, M. & Zimmermann, J. (1995). Purification, characterization and specificity of chondroitin lyases and glycuronidase from *Flavobacterium heparinum*. *Biochem. J.* **312**, 569–577.
26. Jedrzejewski, M. J. (2004). Extracellular virulence factors of *Streptococcus pneumoniae*. *Front Biosci.* **9**, 891–914.
27. Scott, J. E. & Thomlinson, A. M. (1998). The structure of interfibrillar proteoglycan bridges (shape modules) in extracellular matrix of fibrous connective tissues and their stability in various chemical environments. *J. Anat.* **192**, 391–405.
28. Turner, R. E., Lin, P. Y. & Cowman, M. K. (1988). Self-association of hyaluronate segments in aqueous NaCl solution. *Arch. Biochem. Biophys.* **265**, 484–495.
29. McPherson, A. (1999). *Crystallization of Biological Macromolecules*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
30. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
31. Vagin, A. A. & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Crystallog.* **30**, 1022–1025.
32. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
33. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
34. Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458–463.

35. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
36. Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallog. sect. A*, **42**, 140–149.
37. Brunger, A. T. (1992). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
38. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
39. Kleywegt, G. J. (1999). Experimental assessment of differences between related protein crystal structures. *Acta Crystallog. sect. D*, **55**, 1878–1884.
40. Carson, M. (1997). Ribbons. *Methods Enzymol.* **277**, 493–505.
41. Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.
42. Lindahl, E., Hess, B. & van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **7**, 306–317.
43. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236.
44. Kony, D., Damm, W., Stoll, S. & Van Gunsteren, W. F. (2002). An improved OPLS-AA force field for carbohydrates. *J. Comput. Chem.* **23**, 1416–1429.
45. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. (1997). LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472.
46. Berendsen, H. J. C., Postma, J. P. M., DiNola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
47. Darden, T., York, D. & Pedersen, L. (1993). Particle mesh Ewald: an  $N \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092.

*Edited by M. Guss*

(Received 6 December 2005; received in revised form 20 February 2006; accepted 23 February 2006)  
Available online 13 March 2006