



**HAL**  
open science

## Functional and structural insights into human N-deacetylase/N-sulfotransferase activities

Sylvain D Vallet, Thibault Annaval, Romain R Vives, Emeline Richard,  
Jerome Hénault, Christine Le Narvor, David Bonnaffé, Bernard Priem,  
Rebekka Wild, Hugues Lortat-jacob

### ► To cite this version:

Sylvain D Vallet, Thibault Annaval, Romain R Vives, Emeline Richard, Jerome Hénault, et al.. Functional and structural insights into human N-deacetylase/N-sulfotransferase activities. *Proteoglycan Research*, 2023, 1 (3), pp.e8. 10.1002/pgr2.8 . hal-04214217

**HAL Id: hal-04214217**


**<https://hal.science/hal-04214217v1>**

Submitted on 21 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Functional and structural insights into human *N*-deacetylase/*N*-sulfotransferase activities

Sylvain D. Vallet<sup>1</sup> | Thibault Annaval<sup>1</sup> | Romain R. Vives<sup>1</sup> | Emeline Richard<sup>2</sup> |  
 Jerome Hénault<sup>3</sup> | Christine Le Narvor<sup>3</sup> | David Bonnaffé<sup>3</sup> | Bernard Priem<sup>2</sup> |  
 Rebekka Wild<sup>1</sup> | Hugues Lortat-Jacob<sup>1</sup> 

<sup>1</sup>Université Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale, Grenoble, France

<sup>2</sup>Université Grenoble Alpes, CNRS, Centre de Recherches sur les Macromolécules Végétales, Grenoble, France

<sup>3</sup>Université Paris-Saclay, CNRS, Institut de chimie moléculaire et des matériaux d'Orsay, Orsay, France

## Correspondence

Hugues Lortat-Jacob.

Email: [hugues.lortat-jacob@ibs.fr](mailto:hugues.lortat-jacob@ibs.fr)

Rebekka Wild.

Email: [rebekka.wild@ibs.fr](mailto:rebekka.wild@ibs.fr)

## Funding information

The Agence Nationale de la Recherche, Grant/Award Numbers: ANR-18-CE11-0006-01, ANR-21-CE29-0022; The "Investissements d'avenir" program Glyco@Alps, Grant/Award Number: ANR-15-IDEX-02; The CNRS ATIP-Avenir program; FRISBI, Grant/Award Number: ANR-10-INBS-0005-02; GRAL, Grant/Award Number: ANR-17-EURE-0003

## Abstract

Heparan sulfate (HS) is a linear polysaccharide composed of a glucuronic acid (GlcA)-*N*-acetyl-glucosamine (GlcNAc) disaccharide repeat motif, polymerized by the EXT1-EXT2 complex. It is extensively modified by a series of Golgi localized enzymes, that generate distinct saccharide sequences involved in the binding and the regulation of numerous protein partners. *N*-deacetylase/*N*-sulfotransferase (NDST), of which four isoforms have been identified in mammals, are involved in the first step of this process and catalyze both the *N*-deacetylation of the GlcNAc residues into GlcNH<sub>2</sub> and its re-*N*-sulfation into GlcNS residues. Further modifications of the HS chain depend on this first maturation event, NDST action is, therefore, key to HS biosynthesis. However, although the sulfotransferase domain of NDST1 has been characterized at the structural level some 20 years ago, information on the overall structure and activity of the enzyme are still lacking. Here, we report the characterization of the two most expressed NDSTs in humans, NDST1 and NDST2, and a model structure of NDST1 homodimer using cryoelectron microscopy combined with AlphaFold2 modeling. Structure-driven mutagenesis along with two bioassays to follow the protein activities allowed us to characterize the kinetics of the deacetylation and sulfoaddition and to identify the residue H529 as necessary for *N*-deacetylation. These results shed light on a poorly understood family of enzymes and will help deciphering the molecular basis for HS and heparin maturation.

## KEYWORDS

biosynthesis, glycobiology, glycosaminoglycan, heparan sulfate, heparin, NDST

**Abbreviations:** AF2, AlphaFold2; AST IV, aryl sulfotransferase IV; BLI, bio-layer interferometry; dp, degree of polymerization; GlcNH<sub>2</sub>, glucosamine; GlcNS, *N*-sulfated glucosamine; HP, heparin; HS, heparan sulfate; nano-DSF, nano-differential scanning fluorimetry; NDAC, *N*-deacetylase; NDST, *N*-deacetylase/*N*-sulfotransferase; NSt, *N*-sulfotransferase; PAP, adenosine-3',5'-bisphosphate; PNPS, potassium 4-nitrophenyl sulfate; RPIP-HPLC, reversed-phase ion-pairing high-performance liquid chromatography; SEC, size-exclusion chromatography; SPR, surface plasmon resonance; TEV, tobacco etch virus.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Proteoglycan Research* published by Wiley Periodicals LLC.

## INTRODUCTION

Heparan sulfate (HS) is present in every animal tissue both at the cell surface and within the extracellular matrix. There, it regulates the local concentration, stability, conformation, and biological activity of numerous protein ligands, including most cytokines, morphogens, growth factors and their receptors, enzymes, cellular adhesion, and extracellular matrix proteins as well as many proteins involved in host–pathogen interactions.<sup>1</sup> HS thus fulfills a particularly wide range of functions, during both developmental and physiological processes, including cell adhesion, migration, proliferation and differentiation, cellular signaling, extracellular matrix assembly, but also during numerous pathological disorders, such as cancer or infectious and neurodegenerative diseases.<sup>2–8</sup>

A wealth of studies showed that protein binding to HS relies on specific saccharide domains, patterned along the polymer backbone during biosynthesis,<sup>9–13</sup> but the mechanism by which these saccharide domains are assembled remains poorly understood. The HS polymer is first elongated by the EXT1–EXT2 complex that catalyzes the alternative addition of a glucuronic acid (GlcA) and an *N*-acetylglucosamine (GlcNAc) molecule<sup>14–16</sup> and then structurally diversified by a set of Golgi-localized enzymes. The bifunctional *N*-deacetylase/*N*-sulfotransferases (NDSTs) first catalyze *N*-deacetylation and re-*N*-sulfation of GlcNAc, giving rise to glucosamine (GlcNH<sub>2</sub>) and *N*-sulfated glucosamine (GlcNS) residues, respectively. The C5-epimerase then converts GlcNS-adjacent GlcA into iduronic acid (IdoA) and finally various *O*-sulfations can be added at the C6 and C3 positions of the GlcNAc/GlcNS and at the C2 positions of the IdoA or GlcA residues.<sup>17,18</sup>

NDSTs play a major role in generating specific HS motifs. Previous studies revealed that the generated GlcNS moieties are not randomly distributed along the chain but mostly occur in contiguous stretches of usually three to six disaccharides, suggesting that NDSTs work in a processive manner.<sup>19,20</sup> The presence of GlcNS residues, in turn, locally enhances additional subsequent modifications. NDST action thus defines the position and pattern of the protein binding motifs within the polymer chain.<sup>21</sup>

In vertebrates, there are four isoforms of NDSTs. NDST1 and NDST2 are expressed in all embryonic and adult tissues, at albeit different levels, while NDST3 and NDST4 are primarily expressed during development and, at low level, in the adult brain. The enzymes display different ratio of NDST activity that, in addition to their tissue distribution, could account for the observed structural diversity of HS.<sup>22,23</sup> Major efforts have been made to characterize both the NDSTs' *N*-deacetylase and the *N*-sulfotransferase activities (named thereafter NDAc and NSt). Current methodologies are mostly based on the quantification of [<sup>3</sup>H] acetate released from [<sup>3</sup>H] acetyl labeled GlcA–GlcNAc polymer and the incorporation of [<sup>35</sup>S], from [<sup>35</sup>S]-labeled PAPS used as sulfate donor,<sup>24</sup> or the identification of the unsubstituted GlcNH<sub>2</sub> residues by a monoclonal antibody.<sup>25</sup> Further progress, however, has been hampered by the lack of an easy-to-handle quantitative methodology to follow biological activities and challenges in generating highly pure recombinant enzymes.

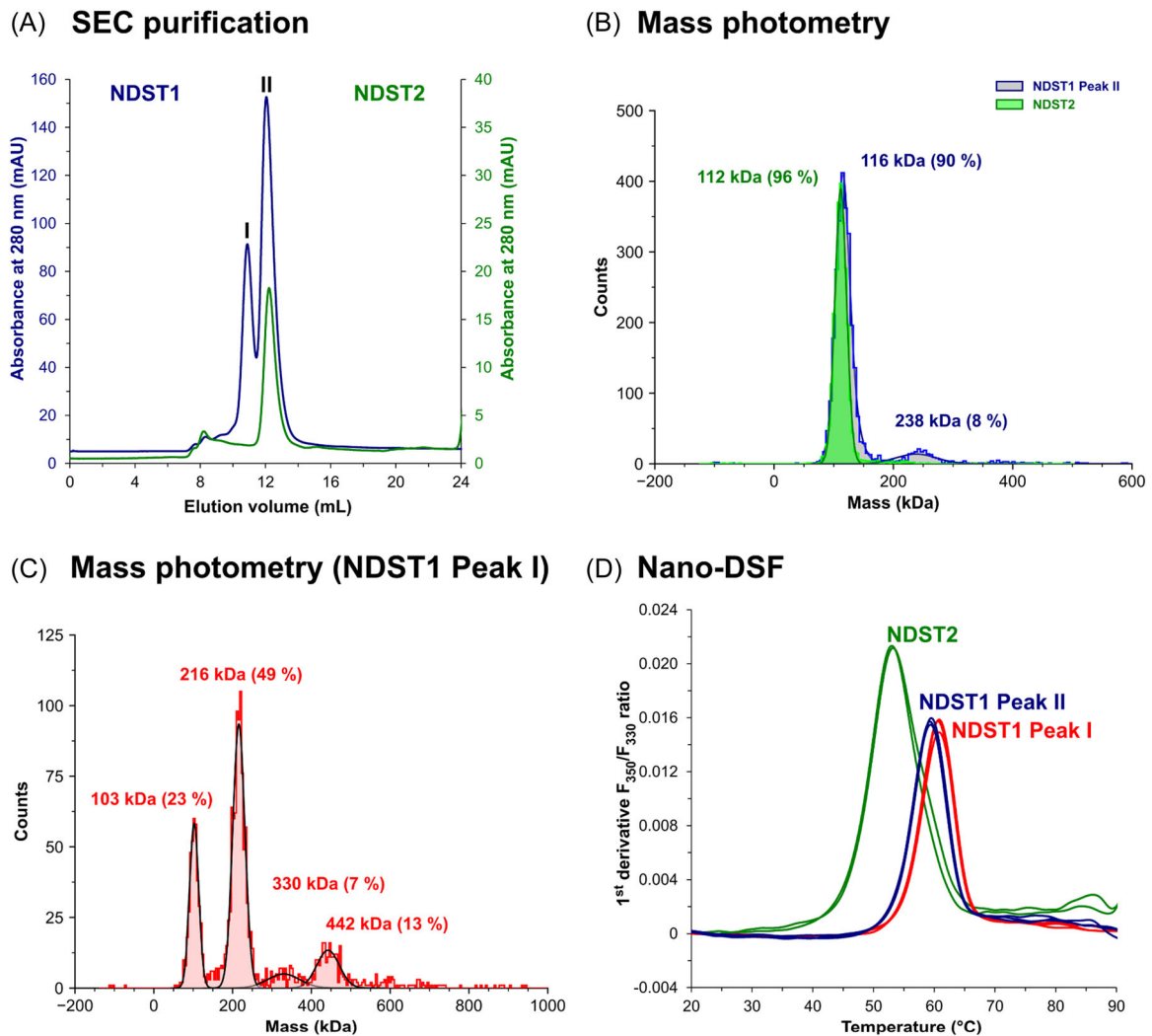
In the present work, we have established a procedure for the recombinant expression and purification of NDSTs and set up two new assays to follow NDAc and NSt activities independently of each other. Using cryoelectron microscopy (cryo-EM), we have determined the structure of an NDST1 homodimeric complex at intermediate resolution, which in combination with an AlphaFold2 prediction, revealed the overall architecture of the bifunctional enzyme as well as the localization of the NDAc active site and its position with regard to the NSt active site.

## RESULTS

### Expression and purification of NDST1 and NDST2

Initial attempts to express NDST1 and NDST2 in human endothelial kidney 293F cells (HEK 293F) were unsuccessful as the proteins were not secreted in the extracellular medium, although constructs were lacking the N-terminal transmembrane helices. The proteins were thus fused to the C-terminus of the Secreted Embryonic Alkaline Phosphatase (SEAP-NDST-His) to improve expression and secretion. HEK 293F cells were transiently transfected and secreted proteins were purified from the conditioned medium 4 days posttransfection using immobilized nickel affinity chromatography (IMAC) (Supporting Information: Figure 1). Fusion proteins were then digested by tobacco etch virus (TEV) protease to remove the SEAP and NDSTs were further purified with a second IMAC step followed by size-exclusion chromatography (SEC) and concentrated up to 3.5 mg/mL.

Expression yields were two- to threefold higher for NDST1 than for NDST2. The size-SEC elution profile of NDSTs featured two major peaks for NDST1 and a single peak for NDST2 (Figure 1A). These peaks were further characterized by mass photometry (Figure 1B,C). In NDST1 Peak II and NDST2 single peak (Figure 1B), both proteins were found to be mostly in a monomeric form with a molecular mass of ~110 kDa (>90%), likely corresponding to the His-tagged enzymes (98 kDa) plus their glycosylations (~6 kDa apparent shift on SDS-PAGE after Peptide *N*-glycosidase F digestion for NDST1; Supporting Information: Figure 2). This is in agreement with previously published data identifying two glycosylation sites in the mouse ortholog (N231 and N351 in human NDST1<sup>26</sup>), while two additional sites, N401 and N667, are predicted to be glycosylated based on the NDST1 sequence. NDST2 sequence contains six potential *N*-glycosylation sites, including one that was identified in the mouse ortholog (N802 in human NDST2 sequence<sup>26</sup>). Mass photometry analysis of NDST1 showed that Peak I comprised a mixture of particles of different sizes (Figure 1C), with molecular masses expected for a monomer (23%), dimer (49%), and tetramer (13%). Addition of 3'-phosphoadenosine-5'-phosphate/3'-phosphoadenosine-5'-phosphosulfate to peak I NDST1 sample did not modify the amount of monomer observed by mass photometry, which still contained ~20% of particles (data not shown). Finally, the stability of NDST1 and NDST2 was assessed by nano-differential scanning fluorimetry (nano-DSF) in presence of 10 mM manganese, a



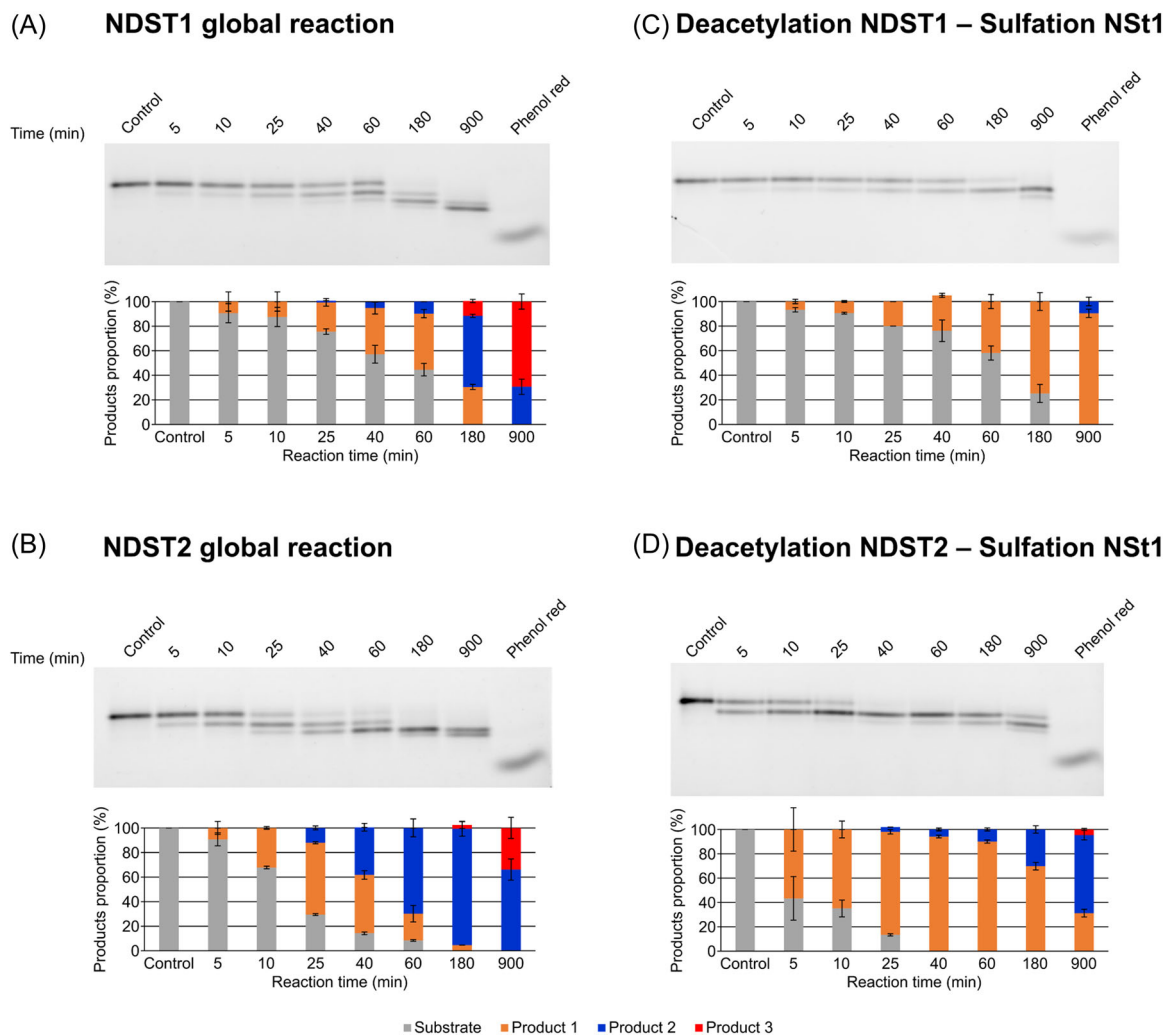
**FIGURE 1** Biochemical characterization of human recombinant NDST1 and NDST2. Both proteins were subjected to size-exclusion chromatography in 50 mM MES, 150 mM NaCl, pH 6.5 on a superdex 200 increase column (Cytiva) as a last purification step (A). Oligomeric states of proteins were analyzed by mass photometry at a protein concentration of 25–50 nM (B, C). The proportion of each species and the calculated molecular weights are shown next to each peak. Thermal stability of NDST1 (Peak I and Peak II, 0.3 mg/mL) and NDST2 (0.16 mg/mL) was assessed by nanodifferential scanning fluorimetry in the presence of 10 mM  $\text{MnCl}_2$  (D). The first derivative of the ratio of fluorescence emission at 350 and 330 nm is shown ( $n = 2$  for NDST2,  $n = 3$  for NDST1 Peak I or Peak II). DSF, differential scanning fluorimetry; NDST, *N*-deacetylase/*N*-sulfotransferase; NDST1, sulfotransferase domain 1; SEC, size-exclusion chromatography.

necessary cofactor for enzyme activity (Figure 1D). NDST1 monomeric form had a measured  $T_m$  of 59.4°C, whereas NDST2 was less stable, with a melting temperature of 53.2°C. NDST1 “peak I” (which mostly comprises oligomeric forms of the enzyme) had a  $T_m$  of 60.7°C, indicating that oligomerization does not increase the stability.

### Kinetics analysis of NDST1 and NDST2

To follow the enzymes' activities, we first developed a fluorophore-assisted carbohydrate electrophoresis (FACE) approach, using a chemically synthesized octasaccharide substrate [(GlcA-GlcNAc)<sub>4</sub>, degree of polymerization 8, dp8] bearing an Alexa Fluor 430 (AF<sub>430</sub>) moiety at its reducing end. During a time course experiment, in which

NDST1 or NDST2 were incubated with dp8-AF<sub>430</sub> and a sulfate donor/regenerating system composed of adenosine-3',5'-bisphosphate (PAP), potassium 4-nitrophenyl sulfate, (PNPS), and aryl sulfotransferase IV (AST IV), three additional bands were observed and quantified by densitometry (Figure 2A,B). We thus concluded that three of the four GlcNAc were *N*-deacetylated and re-*N*-sulfated by the enzymes. The last GlcNAc residue, on which the AF<sub>430</sub> is linked, might not be accessible possibly because of the linkage itself or the absence of the following GlcA residue. This analysis further showed kinetic differences between the two enzymes, with NDST2 being more efficient than NDST1, especially during the first 60 min of enzyme reaction. In the absence of sulfate donor, no change in migration was observed for the dp8-AF<sub>430</sub> upon incubation with the enzymes, suggesting that the *N*-deacetylation



**FIGURE 2** Comparison of NDST1 and NDST2 activities on a fluorescently labeled dp8-AF<sub>430</sub> substrate by FACE. To study the global reactions, 0.4  $\mu$ M of NDST1 (A) or NDST2 (B) and 5  $\mu$ M Alexa-dp8 were incubated with MnCl<sub>2</sub>, PAP, PNPS, and AST IV for up to 900 min. To follow the *N*-deacetylation step, dp8-AF<sub>430</sub> was incubated with NDST1 (C) or NDST2 (D) in presence of MnCl<sub>2</sub> only, for up to 900 min, followed by sulfation by NSt1 in the presence of PAP, PNPS, and AST IV for 15 h. Reaction products were separated on a 25% polyacrylamide gel, visualized and quantified using a ChemiDoc MP system and Image Lab 6.1 software (BioRad). The control lane corresponds to the dp8-AF<sub>430</sub> substrate incubated for 900 min without NDST1 or NDST2. A representative gel is shown and histogram data are averaged from two independent reactions. AF<sub>430</sub>, Alexa Fluor 430; AST IV, aryl sulfotransferase IV; dp8, degree of polymerization 8; FACE, fluorophore-assisted carbohydrate electrophoresis; NDST, *N*-deacetylase/*N*-sulfotransferase; NSt1, sulfotransferase domain 1; PAP, adenosine-3',5'-bisphosphate; PNPS, potassium 4-nitrophenyl sulfate.

alone did not modify the electrophoretic mobility of the substrate (Supporting Information: Figure 3). Hence, to monitor the *N*-deacetylation step on its own, dp8-AF<sub>430</sub> was incubated with NDST1 or NDST2 in the presence of MnCl<sub>2</sub> only. Samples at different time points were heat inactivated and the generated products were treated with the sulfotransferase domain of NDST1 (NSt1) in the presence of the sulfate donor/regenerating system. For NDST1, the appearance of a single additional band was observed, until 900 min (~7% of conversion), indicating that a single GlcNAc was initially deacetylated in the absence of the sulfation reaction, a process that took more than 3 h to be completed (Figure 2C). A second faint band was visible, only after 900 min incubation. This indicates that NDST1 NDAc activity is no longer processive and supports the view that

NDST1 works efficiently only when the NDAc and NSt reactions are coupled (see also Supporting Information: Figure 3). Such analysis has never been performed for NDST2. With this isoform, the first deacetylation reaction appeared after 5 min only (~57% of conversion) and was completed before 40 min, while a second and a third additional band appeared after 1 h (Figure 2D). This indicates that this isoform, in contrast to NDST1, remains able to deacetylate several GlcNAc in the absence of any sulfate donor, and at a much higher rate. It is worth noting that the same results were obtained when resulfation was performed using an NDST1 single point mutant that cannot catalyze deacetylation (see below), which rules out the possibility that the substrate was fully deacetylated and NSt1 was only able to resulfate at one position.

To further quantify the enzymes activities and investigate the deacetylation and resulfation processes kinetically, we next devised a second approach in which NDSTs were incubated for 1 h, with increasing concentrations of (GlcA-GlcNAc)<sub>n</sub> polymer (170 kDa, Supporting Information: Figure 4), either in the presence or the absence of a sulfate donor/regenerating system, to follow the global or the *N*-deacetylation activity, respectively. In the latter case, the resulting GlcNH<sub>2</sub> residues were resulfated with NSt1, as described above. To follow the *N*-sulfation process, NDSTs, along with the sulfate donor/regenerating system, were incubated with (GlcA-GlcNH<sub>2</sub>)<sub>n</sub> polymer, which was prepared by chemical *N*-deacetylation of heparosan (Supporting Information: Figure 5).

The products of these reactions were next extensively depolymerized by heparinase III treatment, after which the generated disaccharides were quantified by high-performance liquid chromatography (HPLC) and enzymatic rates were fitted using the Michaelis–Menten equation (Supporting Information: Figure 6) to determine the kinetics parameters (Table 1). NDST2 featured a higher apparent affinity (lower *K<sub>M</sub>*) for heparosan than NDST1, that likely results from the higher affinity of its NDAc domain for the substrate, while the apparent affinities of the two enzymes for the *N*-deacetylated substrate were similar. The NDAc catalytic activity of NDST2 was higher than that of NDST1, in agreement with the FACE analysis, the opposite being observed for the NSt activities. These results indicate that the *N*-deacetylation step is rate limiting for NDST1, while the re-*N*-sulfation step is limiting for NDST2.

### Binding kinetics analysis of NDST1 and 2 with heparosan

To analyze the affinity of NDST1 and NDST2 for heparosan more directly, we next used a surface plasmon resonance binding assay in which His-tagged NDSTs were captured on Ni-NTA activated sensorchips. Typical binding curves were obtained when increasing concentrations of a 39 kDa heparosan (Supporting Information:

**TABLE 1** NDST1 and NDST2 kinetics parameters for conversion of heparosan (NDAc and complete reactions) and *N*-deacetylated heparosan (NSt reaction) determined following Michaelis–Menten equation (*n* = 2).

Enzyme	Reaction	<i>K<sub>M</sub></i> (μM)	<i>k<sub>cat</sub></i> (min <sup>-1</sup> )
NDST1	NDAc	16.42 ± 1.42	0.239 ± 0.008
	NSt	35.88 ± 7.4	0.975 ± 0.089
	Complete	7.29 ± 1.74	0.23 ± 0.001
NDST2	NDAc	12.04 ± 0.2	0.476 ± 0.008
	NSt	36.33 ± 4.61	0.362 ± 0.023
	Complete	4.07 ± 2.27	0.155 ± 0.002

Note: Enzymatic rates and representations used are shown in Supporting Information: Figure 6.

Abbreviations: NDAc, *N*-deacetylase; NSt, *N*-sulfotransferase.

Figure 7) were flowed above the surfaces (Supporting Information: Figure 8). Affinity was determined using steady-state analysis and showed that NDSTs feature low affinity for heparosan (*K<sub>D</sub>*: 20.5 and 563 μM for NDST1 and NDST2, respectively).

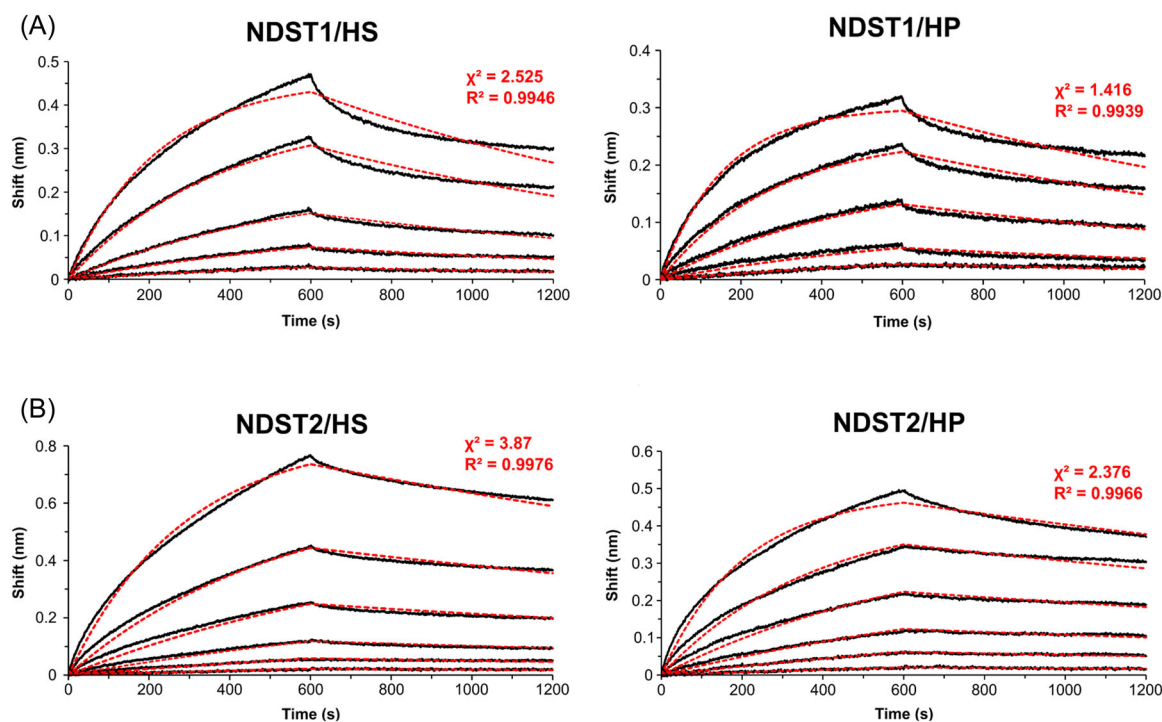
### Binding kinetics analysis of NDST1 and 2 with HS/heparin (HP)

*N*-sulfated sequences, present in HS and HP are known inhibitors of both NDST1 and NDST2.<sup>25</sup> Thus, NDSTs interaction with HS and HP were also analyzed using Bio-Layer Interferometry (BLI), in which each glycosaminoglycan was captured on Streptavidin biosensors (SA, Sartorius). Both proteins strongly interacted with HP and HS with affinities in the low nanomolar range (Figure 3, Table 2), contributed in particular by high stability of the formed complexes (half-live of complexes: NDST1/HS = 13 min, NDST1/HP = 18 min, NDST2/HS = 25 min, NDST2/HP = 40 min). NDST1 and NDST2 bound with higher affinity to HP than HS (two- to three-fold difference in *K<sub>D</sub>* values) and NDST1 showed higher affinity to both GAGs than NDST2. Such a high affinity and complex stability compared to the interaction of NDSTs with heparosan support the view that these molecules are effective inhibitors of both NDSTs.<sup>25</sup> To further investigate this possibility, the deacetylase and global activities of NDST1 and NDST2 in the presence of HP were measured, after 1 h, using the FACE assay. We observed that HP strongly inhibited the deacetylation step catalyzed by NDST1 (89.4% and, to a lesser extent, that catalyzed by NDST2 (51.5%), when used at 10 μM. The overall activities of the enzymes were also inhibited, but less efficiently (59.7% and 18.7% for NDST1 and NDST2 respectively for 10 μM of HP), suggesting that affinity of the enzymes for the substrate increased when a first substrate residue has been sulfated (Supporting Information: Figure 9).

### Architecture of the NDST1 homodimer

SEC and mass photometry experiments revealed that recombinant NDST1 forms oligomers in solution (Figure 1). This prompted us to characterize NDST1 assemblies in more detail using single-particle cryo-EM. Cryo-EM grids containing NDST1 protein from SEC peak I were prepared. Recorded micrographs suggested that the sample was highly heterogeneous, and several rounds of two-dimensional (2D) classification were performed to sort particles in monomeric, dimeric, tetrameric, and higher-oligomeric states (Supporting Information: Figure 10). A majority of recombinant NDST1 was found to form dimers, in good agreement with mass photometry experiments (Figure 1C). We decided to pursue structure determination focusing on the NDST1 dimer.

Initial attempts to generate an ab initio three-dimensional (3D) model of the dimer were unsuccessful, likely due to a strong preferred particle orientation on the grid (Supporting Information: Figure 10). To overcome this issue, we collected a data set using a



**FIGURE 3** Kinetics analysis of the interactions between HS or HP with NDST1 (A) and NDST2 (B) by Bio-Layer Interferometry. Each sensorgram was obtained with NDST1 at 1.56, 3.13, 6.25, 12.5, and 25 nM (A), and with NDST2 at 3.13, 6.25, 12.5, 25, 50, and 100 nM (B), from bottom to top. Experimental data (colored in black) were fitted to the 1:1 interaction model (red dotted line).  $\chi^2$  and  $R^2$  values are indicated in red, next to the curves. HP, heparin; HS, heparan sulfate; NDST, N-deacetylase/N-sulfotransferase.

**TABLE 2** Kinetics and affinity parameters of the interactions between NDST1 or NDST2 and HS and HP ( $n = 2$ ).

Analyte/ ligand	$K_D$ (nM)	$k_a$ (1/Ms)	$k_d$ (1/s)
NDST1			
HS	$5.36 \pm 0.14$	$(1.65 \pm 0.13) \times 10^5$	$(8.86 \pm 0.94) \times 10^{-4}$
HP	$2.93 \pm 0$	$(2.19 \pm 0.11) \times 10^5$	$(6.42 \pm 0.32) \times 10^{-4}$
NDST2			
HS	$17.5 \pm 5.7$	$(2.77 \pm 0.36) \times 10^4$	$(4.64 \pm 0.95) \times 10^{-4}$
HP	$5.21 \pm 0.2$	$(5.56 \pm 0.57) \times 10^4$	$(2.9 \pm 0.14) \times 10^{-4}$

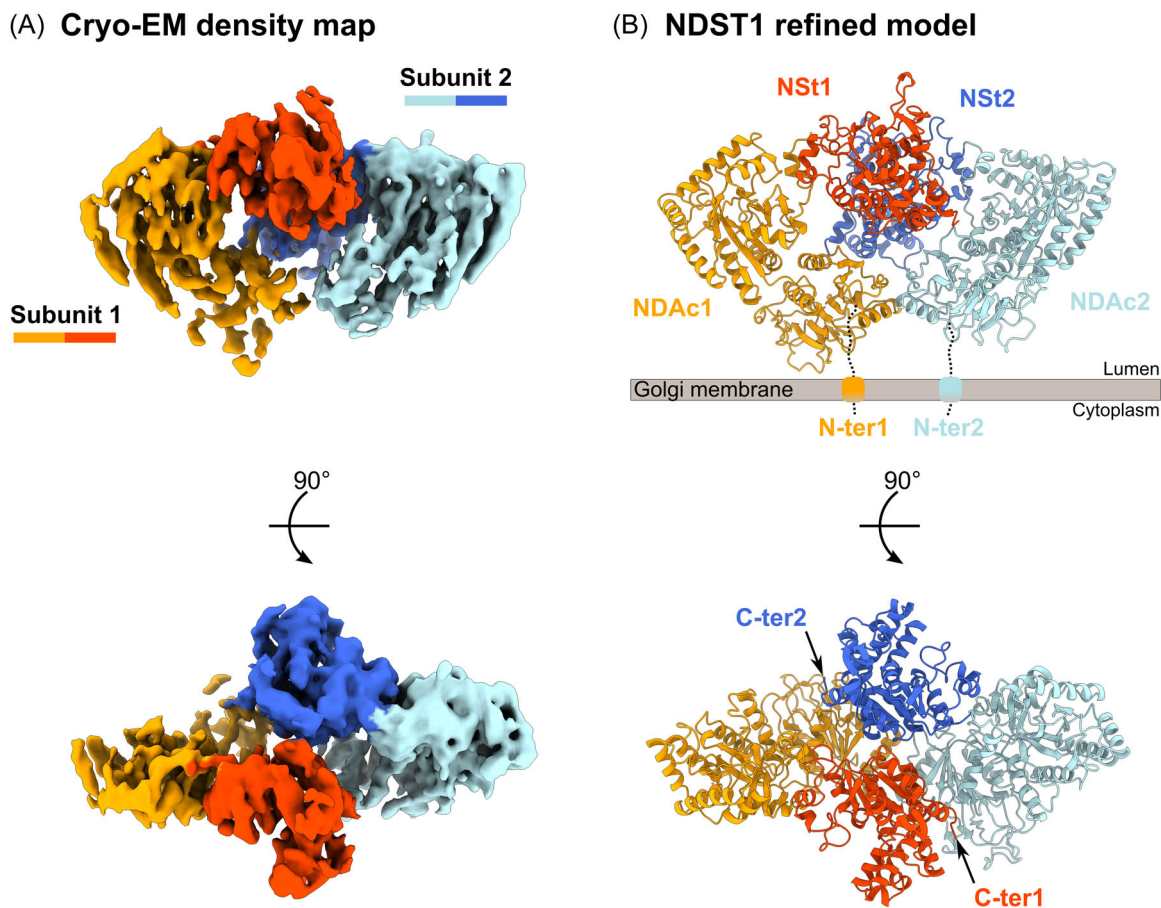
Abbreviations: HP, heparin; HS, heparan sulfate; NDST, N-deacetylase/N-sulfotransferase.

30° stage tilt and resulting 2D class averages confirmed the presence of novel particle orientations compared to the nontilted data set (Supporting Information: Figure 11). The tilted data helped to overcome initial difficulties to reconstruct an ab initio 3D model. We generated 3D reconstructions using either both data sets (tilted and nontilted) or only the tilted micrographs. The latter resulted into a 3D reconstruction of the NDST1 homodimer at a nominal resolution of 4.5 Å with less streaking artifacts compared to the structure using combined data sets (Figure 4A, Supporting Information: Figures 11 and 12). The angular distribution plot of particles still showed a preferred orientation but with an extended cone

(Supporting Information: Figure 11). The obtained cryo-EM map was, however, not of sufficient resolution to build a model of the NDST1 homodimer de novo.

We thus predicted the structure of the NDST1 homodimer (residues 65–882) using AlphaFold2<sup>27,28</sup> and the resulting model was docked into the cryo-EM map. A very good overall fit was obtained with only a few helices and smaller domains slightly outside the density, as well as some loops and the N-terminus, which were missing in the EM map (Supporting Information: Figure 13A). Residues 65–79 and 65–76 of subunit 1 and 2, respectively, were removed manually from the model and model refinement was performed with FlexEM<sup>29</sup> (Figure 4B, Supporting Information: Figures 13B and 14). Some short loops, mostly located in the NDAc domains, remained outside the density map (Supporting Information: Figure 15).

The NDST1 homodimer features a C2 symmetry and adopts an elongated shape, in which the two V-shaped monomers are interlocked. The N-termini of both monomers locate to the same complex surface, likely the side facing the Golgi membrane. The dimer interface accounts for 1195 Å<sup>2</sup> and involves interactions between the N-terminal NDAc domain of one monomer with both, the NDAc domain (544 Å<sup>2</sup>) and NSt domain (164 and 322 Å<sup>2</sup>) of the other monomer. Interactions between NSt domains (165 Å<sup>2</sup>) were also observed. Analysis of contact residues at the interface showed that they are localized in the 90–162 and 769–881 regions, which correspond to the N-terminal portion of the NDAc domain and the



**FIGURE 4** Cryo-EM structure of the NDST1 homodimer. (A) Cryo-EM density map, with NDST1 subunit 1 colored in orange/red and subunit 2 in blue. Upper panel shows the complex from the side, bottom view is rotated at 90°. (B) Structure of the NDST1 homodimer shown in ribbon representation. An AlphaFold2-predicted model was docked and refined using FlexEM. Cryo-EM, cryoelectron microscopy; NDAc1/2, N-deacetylase domain 1/2; NDST, N-deacetylase/N-sulfotransferase; NSt1/2, N-sulfotransferase domain 1/2.

C-terminal portion of the NSt domain. Of note the AlphaFold2 predicted monomer of NDST1 superimposed well with subunit 1 from the refined dimer structure (Supporting Information: Figure 16A), suggesting that NDST1 monomers do not undergo any major conformational changes upon complex formation. This is further supported by the observation that the dimer subunit 1 fits well into a low-resolution cryo-EM map of the monomer (Supporting Information: Figure 16B), and a good fit of the AlphaFold2 monomer model as well as the low-resolution monomer map into the cryo-EM map of the dimer (Supporting Information: Figure 16C,D).

### Identification of the NDAc active site

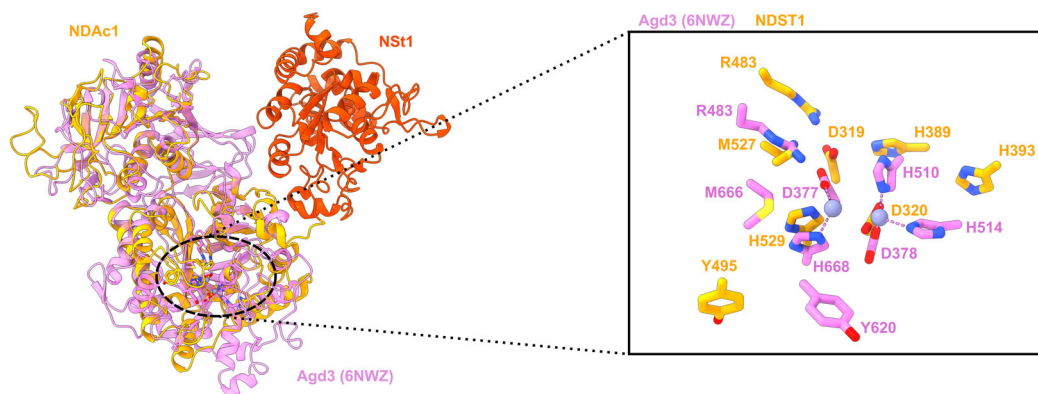
The cryo-EM studies, together with the AlphaFold2 predicted structures of NDST1 (Figure 5A) and NDST2, 3 and 4 (Supporting Information: Figure 17) show that NDSTs consist of three domains, arranged in a V-shaped architecture. Two of them (residues 40–598 in human NDST1), supposedly harbor the active site carrying out N-deacetylation, the C-terminal domain (residues 599–882 in human NDST1) harboring the sulfotransferase activity.<sup>30</sup>

To identify the NDAc active site, we first performed a search for similar structures using DALI.<sup>31</sup> Closest hit was Agd3, an exopolysaccharide deacetylase from *Aspergillus fumigatus* (PDB-ID 6NWZ<sup>32</sup>) that deacetylates  $\alpha$ -1,4-GalNAc oligosaccharides. The Agd3 crystal structure superimposes well with both the N-terminal and central domain of NDST1 with an root-mean-square deviation of 1.12 Å over 110 Ca atoms (Figure 5A).

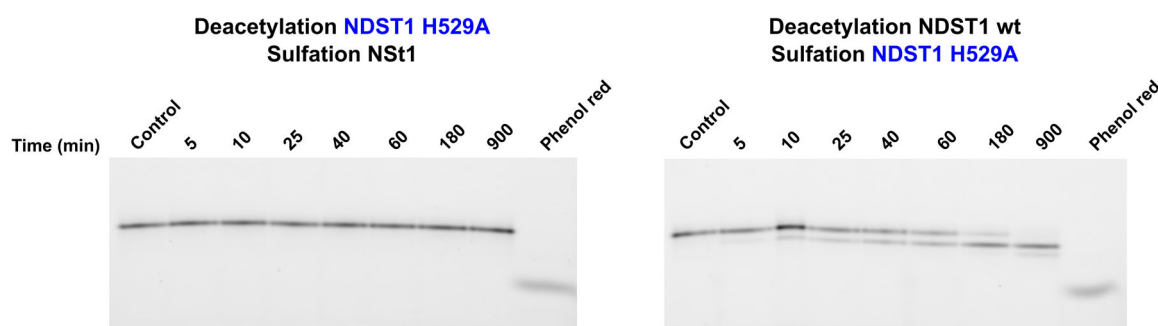
Next, we checked if residues in Agd3 involved in catalyzing N-deacetylation are conserved in NDST1. Importantly, Agd3 requires a divalent cation for its activity and although the Agd3 crystal structure contained a Zn<sup>2+</sup>, it was shown to have highest activity in presence of Mn<sup>2+</sup>, the natural cofactor of NDST1.<sup>33</sup> The metal ion triad responsible for Zn<sup>2+</sup> coordination in Agd3 (D378, H510, and H514) was found to be conserved in NDST1 (D320, H389, and H393) (Figure 5B). In addition, the putative catalytic base and catalytic acid (H668 and D377, respectively) in Agd3 superimposed well with residues H529 and D319 in NDST1. M666 and Y620, which were postulated to be involved in the catalytic site formation or the substrate alignment and the substrate binding, respectively, in Agd3<sup>32</sup> were also conserved in human NDST1 (M527 and Y495), even if the latter was further away from the active site in NDST1 structure. The strong conservation of the spatial arrangement of



## (A) Agd3 superimposition on NDST1 monomer



## (B) NDST1 H529A FACE assays



**FIGURE 5** Characterization of the *N*-deacetylase active site. (A) Superimposition of Agd3 deacetylase (PDB ID: 6NWZ, in transparent mauve) with NDST1 subunit 1 refined structure (orange), and close-up view on the residues of Agd3 involved in metal ion coordination and catalysis conserved in NDST1 (in sticks). NDST1 and Agd3 residues are indicated in orange and mauve, respectively. (B) Analysis of NDST1 H529A mutant (peak II) activity on a fluorescently labeled dp8 by FACE. *N*-deacetylation step was followed in the presence of 10 mM MnCl<sub>2</sub> without any sulfate donor, then NST1, PAP, PNPS, and AST IV were added to sulfate any *N*-glucosamine residues generated. Following the same protocol as above, NDST1 wt was used for the *N*-deacetylation followed by sulfation by NDST1 H529A mutant instead of NST1. Reaction products were separated on a 25% polyacrylamide gel and visualized using a ChemiDoc MP system (BioRad). AF<sub>430</sub>, Alexa Fluor 430; AST IV, aryl sulfotransferase IV; dp8, degree of polymerization 8; FACE, fluorophore-assisted carbohydrate electrophoresis; NDST, *N*-deacetylase/*N*-sulfotransferase; NST1, *N*-sulfotransferase domain 1; PAP, adenosine-3',5'-bisphosphate; PNPS, potassium 4-nitrophenyl sulfate.

catalytic residues suggests that NDST1 and Agd3 share a similar reaction mechanism, and reveals the location of the catalytic site in the central domain of NDST1. Of note, all these residues (i.e., D319, D320, H389, H393, R483, Y495, M527, and H529) are also conserved in human NDST2-4 sequences. Interestingly, this sequence region was more conserved than the rest of the protein when comparing NDST1 and NDST2 or NDST1 sequences from 86 different species (Supporting Information: Figure 18A,B).

### Mutants with impaired NDAc activity

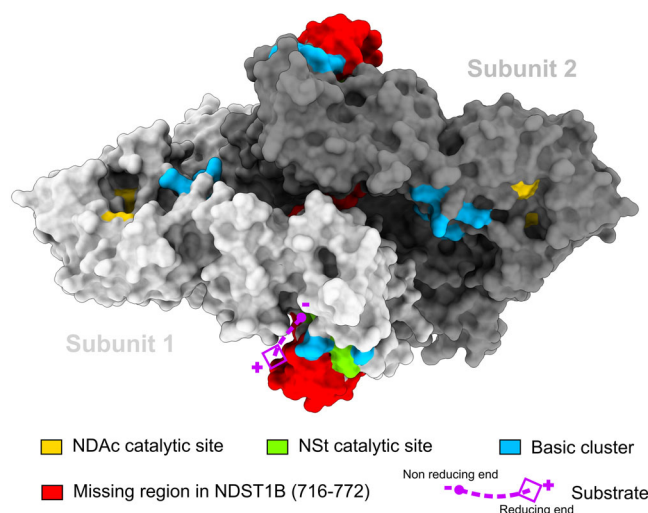
To experimentally define the NDAc active site, we generated NDST1 double (D319N/D320N) and single (H529A) point mutants. NDST1 double mutant D319N/D320N was not secreted by HEK 293F cells upon transfection, suggesting that these residues are important for the stability and/or the folding of the protein. H529A mutant was

expressed in similar amounts as the wt enzyme. SEC, mass photometry, and melting temperature analysis further confirmed that the H529A mutant protein was well folded and was similar to the wt enzyme with regard to molecular mass, oligomerization, and thermal stability (Supporting Information: Figure 19). Consistently with the possibility that the H529A residue could act as a base in the *N*-deacetylation reaction, we observed that this mutant was completely inactive with regard to the *N*-deacetylation reaction on both dp8-AF<sub>430</sub> (Figure 5B) and heparosan substrates, whereas its sulfotransferase activity remained similar to that of the wild-type enzyme (Supporting Information: Figure 20). Derivatization of C486 with *N*-ethylmaleimide was shown to inactivate NDST deacetylase activity.<sup>34</sup> Interestingly, in our model, this residue is close to the identified catalytic site (11 and 13 Å between C486 Cα and H529 or D319 Cα, respectively, Supporting Information: Figure 21), thus possibly supporting the view that *N*-ethylmaleimide hinders access to the active site.

## DISCUSSION

In mammals, the NDST family is composed of four homologous Golgi membrane-anchored enzymes. These bifunctional enzymes contain a catalytic site involved in GlcNAc *N*-deacetylation and another one for re-*N*-sulfation, together assuring the first maturation step of HS. NDSTs act during and after HS polymerization by the EXT1–EXT2 complex and are initiating further modifications of the GAG chain. So far, the molecular mechanism of this first step of HS maturation remained poorly understood. In this study, the ectodomains of NDST1 and NDST2 were purified to homogeneity and functionally and structurally characterized.

NDST1<sup>35</sup> or NDST2<sup>36,37</sup> knock-out mice studies showed that while NDST1 is involved in HS modifications, forming size-limited *N*-sulfated domains, NDST2 rather catalyzes the generation of longer stretches of *N*-sulfated sequences (for review see Kjellén<sup>38</sup>), such as in HP, but also could modulate chain length.<sup>39</sup> To compare the activity of both enzymes, we developed two bioassays that avoid the use of radioactive substrate. We showed that in the absence of the sulfation process, NDST1 displays a limited de-acetylase activity, with a single GlcNAc modified when an octasaccharide was used as a substrate, at a strikingly lower rate than with NDST2 (3 h for NDST1 vs. 40 min for NDST2). In contrast, NDST2 was able to process as much as three GlcNAc. When both deacetylation and sulfation activities were allowed to work simultaneously, up to three GlcNAc were efficiently deacetylated/resulfated. This further supports the view that the two enzymatic reactions are functionally coupled and demonstrate, in particular, that interplay between the two catalytic activities is involved in the processivity and efficacy of the enzymes. In our model, however, the NSt and NDAC sites are 6 nm away from each other and do not appear in line. We could not identify any obvious substrate channel in between the two catalytic sites that could explain how the catalytic rates are enhanced when the two reactions are coupled (see Figure 6 and Supporting Information: Figures 21 and 22, in which the position and orientation of the oligosaccharide substrate was drawn on the NDST's sulfotransferase domain, using the murine 3-*O*-sulfotransferase-1 structure (PDB ID 3UAN) in complex with HS as template<sup>40</sup>). Within the NDST1 homodimer the NDAC1 and NSt2 (or NDAC2 and NSt1) are also 6 nm away from each other, but are more aligned across the dimer than within a single monomer. It is however not clear to ascertain that this could promote catalytic rate enhancement. Rather, it can be assumed that after an initial random encounter, the enzyme remains associated with the substrate and works along the chain, a process that would be more efficient than two random encounters made necessary by the separation of the two domains or the two monomers. Of note, as PAPS concentration was found to directly affect the processive mode of NDST1 and the formation of the *N*-sulfated domain,<sup>19</sup> our assays included a PAPS regenerating system, providing a constant amount of the reagent.



**FIGURE 6** Visualization of the NDAC (gold) and NSt (green) catalytic sites in the refined NDST1 homodimer structure. The three basic clusters contained in each NDST1 monomer are colored in blue and the region missing in NDST1B variant is shown in red. Orientation of the substrate in the NSt catalytic site was identified after superimposition with the murine 3-*O*-sulfotransferase-1 structure (PDB ID 3UAN). NDAC, *N*-deacetylase; NDST, *N*-deacetylase/*N*-sulfotransferase; NSt, *N*-sulfotransferase.

Previous works on NDST1 showed that deacetylase activity plays a key role in the formation of *N*-sulfated domain in HS.<sup>41</sup> Our kinetic analysis demonstrated that the *N*-deacetylation step was rate limiting for NDST1, in agreement with previous study,<sup>20,41</sup> but in contrast, it was the re-*N*-sulfation that was rate limiting for NDST2. The differences we thus report here account for the different degrees of polymerization of the NS domains formed by the action of NDST1 and NDST2 in HS and HP, respectively.

Affinity for heparosan was measured by SPR using the steady-state analysis and returned apparent  $K_D$  in the  $\mu\text{M}$  range (20.5 and 563.5  $\mu\text{M}$  for NDST1 and NDST2, respectively). As 39 kDa heparosan comprises around 100 disaccharide repeats, it might interact with several NDSTs, thus possibly giving rise to an avidity effect, depending on the NDST immobilized density on the surface. The above-indicated values should thus be considered as apparent affinities, much lower than those measured for HS and HP, in the nM range. These differences could explain the inhibitions of NDSTs by *N*-sulfated sequences, present in HS and HP.<sup>25</sup> Interestingly, both enzymes feature basic patches, located at either side of the catalytic site, that could bind HP and HS and thus prevent access of the substrate (Supporting Information: Figures 21 and 22). These high affinities could indicate a role in the regulation of the activity of NDST enzymes. The reaction products HS and HP could therefore, by local sequestration, prevent further heparosan modifications.

SEC, mass photometry, and cryo-EM analysis concurringly suggested that NDST1 forms different oligomeric assemblies in solution. NDST1 has been previously shown by FRET to colocalize with its inactive variant NDST1B, which lacks the region

716–772.<sup>42</sup> This region is found to be distant from the dimeric interface in our structural model (Figure 6). Of note, to the best of our knowledge, we provide for the first time biochemical evidence for a direct physical interaction between members of the NDST family. In the V-shaped monomer, as well as in the homodimer, the NDAc and NSt catalytic sites are ~60 Å away from each other. In the dimer structure, however, the catalytic sites seem to be more aligned (Figure 6), but the functional importance of dimer formation, for example, in promoting processivity of the enzyme, and its significance in vivo, has yet to be investigated and will need further investigations

Using single particle cryo-EM, we were able to obtain a 3D reconstruction of the NDST1 homodimer at an intermediate resolution (4.5 Å). An AlphaFold2 predicted homodimer fitted very well into the experimental EM map. NDST1 adopts a three-domain organization and the same structural organization was observed in AlphaFold2 models of human NDST2-4 (Supporting Information: Figure 17), which is expected considering their high sequence identity (>65%).

No evidence of NDST2 dimerization was observed by either SEC nor mass photometry, despite using the same expression strategy as for NDST1. This was unexpected, because both protein sequences are highly conserved (~73% identity for the Golgi-luminal part). Importantly, most of the residues involved in NDST1 dimer formation (e.g., W150, E153, and K157) are conserved in NDST2 or replaced with amino acids with similar properties (i.e., Lys to Arg). One possible explanation for the absence of NDST2 dimers could be the presence of an N-glycan in the NSt domain of NDST2 (N802). This posttranslational modification was identified in the NDST2 ortholog from mouse and it locates very close to the potential dimer interface.

Superposition of the NDST1 structure with that of the carbohydrate deacetylase Agd3, revealed a high structural similarity and further pointed out the residues H529 and D319 could be important for catalytic activity of NDST1. This was confirmed by H529 substitution with an alanine residue, which resulting NDST1 H529A mutant showed impaired deacetylation activity while its sulfation activity remained unaffected (Figure 5). This is in agreement with the existence of two distinct catalytic sites in the protein.<sup>43</sup>

NDSTs share a three-domain architecture, with the NDAc and the NSt catalytic sites being localized in the middle and the C-terminal domains, respectively (Figure 6). So far, the functional importance of the N-terminal domain remains unknown. Interestingly, the presence of this first domain is also observed for Agd3.<sup>32</sup> It could be involved in complex formation with other Golgi enzymes or be important for protein folding, localization, and stability. Future studies will be needed to unravel potential additional functions of the NDST subdomains, the bias of NDST1 to form oligomers and the functional interplay between the two catalytic activities. Our results may have possible implications for the development of chemoenzymatic approaches aiming at synthesizing HS-like oligosaccharides featuring either limited or extended N-sulfated domains and should permit to investigate the functional relationship between NDSTs and other HS biosynthesis enzymes.

## MATERIALS AND METHODS

### Expression and purification of NDST1 and NDST2

Human NDST1 (P52848, residues 43–882 according to UniProtKB numbering) and NDST2 (P52849, residues 44–883) lacking their N-terminus transmembrane domains were cloned within the pTT22-SEAP vector (National Research Center Canada, Biotechnology Research Institute) in frame with the SEAP (UniProtKB P05187, residues 1–506), and a linker region containing a TEV-cleavage site (NSSSSNNNNNNNNLGENLYFQGAS) on the N-terminal side and a 8-His tag (NNNNNNGHHHHHHHH) at their C-terminus. NDSTs were transiently expressed under a recombinant form in Freestyle 293-F cells (Human Embryonal Kidney 293, HEK 293F, R79007; Invitrogen). Briefly, cells were grown in FreeStyle™ 293 Expression Medium (12338018; Gibco) under agitation at 37°C and 8% CO<sub>2</sub>. When they reached 1.8–2 × 10<sup>6</sup> cells/mL, cells were transfected with polyethylenimine and the SEAP-NDST-His construct. 0.5% (w/v) tryptone (19553; Organotechnie) was added 24 h posttransfection and the medium was collected 3 days later. Cells and cellular debris were removed by two centrifuge steps (500g, 15 min, then 10,000g, 30 min at 4°C) and supernatant was filtered using a pore size of 0.45 µm before purification.

SEAP-NDST1-His and SEAP-NDST2-His were purified by immobilized metal affinity chromatography (HisTrap HP, Cytiva or Protino Ni-NTa Agarose; Macherey-Nagel) in 50 mM Tris, 150 mM NaCl, pH 8, followed by removal of the SEAP fusion protein by TEV-protease cleavage. Finally, proteins were purified by SEC using a Superdex S200 column (HiLoad 16/600, 200 µg or S200 Increase 10/300 GL; Cytiva) pre-equilibrated in 50 mM MES, 150 mM NaCl, pH 6.5. Peak fractions were pooled and concentrated up to 3.5 mg/mL before flash-freezing in liquid nitrogen.

Purification steps were followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis stained by Instant Blue (ISB1L; Sigma-Aldrich).

### Expression and purification of NSt1 domain

The DNA coding for NSt1 (residues 557–882) in fusion with a SUMO and a 6-His tags at its N-terminus was inserted in a pBG102 vector (Center for Structural Biology, Vanderbilt University) and transformed into *Escherichia coli* BL21 bacteria. Bacteria were lysed by sonication in 50 mM Tris, 150 mM NaCl, pH 8. Both tags and imidazole were removed by cleavage during dialysis in the presence of Human Rhinovirus 3C protease. Finally, the dialysate was loaded on a HisTrap HP column (Cytiva) and the flow-through, containing only NSt1 was collected and concentrated up to 33 µM.

## NDST1 mutant generation

NDST1 mutants D319N/D320N and H529A constructs were generated by one-step site-directed mutagenesis using primers listed in Supporting Information: Table 1 and verified by Sanger sequencing.

The resulting proteins were expressed and purified using the same protocol described above for the wild-type protein.

## Preparation of heparosan samples

Heparosan was prepared in recombinant *E. coli* strains expressing various genes involved in K5-heparosan synthesis as followed: Low molecular weight (LMW) heparosan was produced by the strain K-hep in the presence of lactose serving as priming acceptor of polymerization of heparosan as described.<sup>44</sup> This strain was an *E. coli* K-12 derivative carrying recombinant plasmids pBS-*kfiCA* and pBBR-*glcAT-kfiD-kfiB*. Heparosan was produced intracellularly due to the lack of export genes *kps* in *E. coli* K-12. High molecular weight (HMW) heparosan was produced by the strain B-hep carrying recombinant plasmids pBBR-*kfiAB* and pSU-*kfiCD*. This strain was nearly identical to strain HB5 previously used to produce heparosan,<sup>45</sup> except that host strain was an *E. coli* BL21 derivative carrying export genes *kps* thus producing an extracellular heparosan.

Both strains were cultivated in fed-batch conditions on mineral culture medium, as previously described.<sup>44,45</sup> After the first phase of growth on glucose, a feeding solution of glycerol was provided at a flow rate of 7 mL h<sup>-1</sup> L<sup>-1</sup> during 72 h. Then, cultures were centrifuged (8400g, 25 min, room temperature [RT]) to separate the cell pellet from the extracellular medium. LMW heparosan was recovered from pellet of the K-hep culture resuspended in water and heated at 105°C during 40 min. HMW heparosan was recovered from extracellular medium of the B-hep culture. Proteins were discarded by lowering the pH to 3.5 with IR 120 H+ resin and centrifugation (8400g, 25 min, RT) as before. Heparosan was precipitated with ethanol (three volumes, overnight, 4°C) followed by centrifugation (8400g, 25 min, RT), and lyophilization.

## Carbohydrate analyses

Average molecular weight and molecular weight distribution were determined by high-pressure size-exclusion chromatography (HPSEC) (Alliance GPCV 2000; Waters) using an OHPAK SB-G guard column and an OHPAK SB 806 M HQ column (Shodex) with on-line multiangle laser light scattering (MALLS) (DAWN HELEOS II; Wyatt Technology Corp.) fitted with a K5 cell and a laser wavelength of 660 nm, a refractive index detector, and a viscometer. Samples were eluted with 0.1 M NaNO<sub>3</sub> containing 0.03% NaN<sub>3</sub> at 0.5 mL min<sup>-1</sup>. Solvent and samples were filtered through 0.1 and 0.2 μm filter units (Millipore), respectively. Samples were injected at a 2 mg/mL concentration through a 100 μL full loop.

Spectroscopy <sup>13</sup>C and <sup>1</sup>H NMR spectra were recorded with a BRUKER Avance 400 spectrometer operating at a frequency of 100.618 MHz for <sup>13</sup>C and 400.13 MHz for <sup>1</sup>H. Samples were solubilized in D<sub>2</sub>O and analyzed at 353 K. Residual signal of the solvent was used as an internal standard: HOD at 4.25 ppm.

## Activity assays

To follow the *N*-deacetylation step, the *N*-sulfation step or the complete reaction, 0.03, 0.3, 1, 4, 7.5, 10, and 15 mg/mL of HMW heparosan or *N*-deacetylated HMW heparosan were incubated with NDST1 or NDST2 for 1 h at 37°C in a final volume of 25 μL in 50 mM MES, 150 mM NaCl, pH 6.5. Reactions were stopped by heating samples at 100°C and the supernatants fractions were collected and digested by heparinase III overnight in 100 mM sodium acetate, 0.5 mM calcium acetate, pH 7.1.

*N*-deacetylation activity was followed by sulfating unsubstituted glucosamine (GlcNH<sub>3</sub><sup>+</sup>) residues generated by NDST1 or NDST2 because GlcNH<sub>3</sub><sup>+</sup> containing disaccharide elution volume is near the void volume of the column. Briefly, 10 mM MnCl<sub>2</sub> was added in the reaction mix with 0.4 μM NDST1 or 0.1 μM NDST2. After stopping the reaction, 1 μM NSt1 domain was added and sulfation reaction was carried out for 15 h at 37°C in the presence of 0.5 mM adenosine-3',5'-bisphosphate (PAP), 10 mM PNPS, and 27 μM AST IV.

The *N*-sulfation reaction was followed by incubating chemically *N*-deacetylated HMW heparosan with 0.2 μM NDST1 or NDST2 for 1 h at 37°C in presence of 0.5 mM PAP, 10 mM PNPS and 27 μM AST IV. Briefly, HMW heparosan was chemically *N*-deacetylated in the presence of 2 M NaOH and incubation at 56°C for 16 h.<sup>46–48</sup>

The complete reaction was followed by incubating 0.2 and 0.4 μM NDST1 or NDST2 with 10 mM MnCl<sub>2</sub>, 0.5 mM PAP, 10 mM PNPS and 27 μM AST.

Reversed-phase ion-pairing HPLC (RPIP-HPLC) settings used were described in Laguri et al.<sup>49</sup> Briefly, 250 ng of enzymatic products digested by heparinase III were injected on a Luna 5 μm C18 reversed-phase column (4.6 × 300 mm; Phenomenex) in 1.2 mM tetra-*N*-butylammonium hydrogen sulfate, pH 3.15, 8.5% (v/v) acetonitrile at a flow rate of 0.5 mL/min. Oligosaccharides were eluted by a gradient of the running buffer containing 0.6 M NaCl. Postcolumn labeling was conducted in a 1:1 ratio of 0.25% (w/v) 2-cyanoacetamide and 51% (w/v) NaOH at 130°C. Elution was followed with fluorescence (excitation: 346 nm, emission: 410 nm). Conversion percentages were estimated by measuring the proportion of the area of ΔGlcA-GlcNAc (ΔGlcA-GlcNH<sub>3</sub><sup>+</sup> for ST activity) and ΔGlcA-GlcNS peaks.

## Enzymatic modification of synthetic fluorescently-labeled octasaccharide

AF<sub>430</sub> was coupled to the reducing end of a synthetic heparosan octasaccharide (GlcA-GlcNAc)<sub>4</sub> as already described.<sup>16</sup> Reactions were followed by FACE.

Briefly, 5  $\mu\text{M}$  of dp8-AF<sub>430</sub> were incubated with 0.4  $\mu\text{M}$  NDST1 or NDST2 for up to 15 h at 37°C in a final volume of 10  $\mu\text{L}$ . To follow the N-deacetylation reaction, experiments were conducted in the presence of 10 mM MnCl<sub>2</sub>. Reactions were stopped by heating the samples for 5 min at 70°C, then N-deacetylated residues were sulfated by addition of 1  $\mu\text{M}$  NSt1 or NDST1 (wild-type or H529A mutant), 0.5 mM PAP, 10 mM PNPS, and 27.3  $\mu\text{M}$  AST IV for 15 h at 37°C in a final volume of 15  $\mu\text{L}$ . Reactions were stopped by heating at 100°C for 5 min. For the complete reaction, 10 mM MnCl<sub>2</sub>, 0.5 mM PAP, 10 mM PNPS, and 27.3  $\mu\text{M}$  AST IV were included in the one-pot reaction mixture. Finally, reactions were heat-inactivated and 17% (v/v) glycerol was added to the samples before loading on a 25% tris-glycine polyacrylamide gel. For HP-mediated inhibition, the same set up was used but HP (H3393; Sigma-Aldrich) was added at the final concentration of 2.5, 5, or 10  $\mu\text{M}$  in the reaction mixture for 1 h at 37°C. Fluorescent detection was performed using a ChemiDoc MP system (BioRad).

## Thermal stability analysis

Stability of NDST1 wt and mutant proteins stability was assessed by nano-DSF using a Prometheus NT.48 (NanoTemper) instrument with standard capillaries (PR-C002). NDST1 and NDST2 were diluted to 0.3 and 0.16 mg/mL, respectively, in buffer containing 50 mM MES, 150 mM NaCl, pH 6.5, and 10 mM MnCl<sub>2</sub>. Melting curves were recorded while heating samples from 20°C to 95°C with a 2°C/min step and at a power intensity of 40%–80%.

## Mass photometry analysis

Proteins diluted at 25–50 nM in 50 mM MES, 150 mM NaCl, pH 6.5 (MES buffer) were analyzed on OneMP equipment (Refeyn) with high precision cover glasses (24 × 50 mm, 170 ± 5  $\mu\text{m}$ , No 1.5H) and CultureWell reusable gasket (CW-50R-1.0; Grace Bio-Labs). Each measurement was performed after setting the focus with a drop containing 19  $\mu\text{L}$  of MES buffer, followed by addition of 1  $\mu\text{L}$  of the protein solution (0.5–1  $\mu\text{M}$ ). NativeMark™ unstained protein standard (ThermoFisher Scientific) containing proteins of 66, 146, 480, and 1048 kDa diluted in MES buffer was used for calibration. 6000 frames were recorded using a regular field of view (18  $\mu\text{m}^2$ ) and data were analyzed with DiscoverMP v2.5.0 software.

## Cryoelectron data collection and processing

4  $\mu\text{L}$  of NDST1 (SEC, elution peak I) at 0.3 mg/mL in 8 mM MES, 150 mM NaCl, pH 6.5 were applied onto a Quantifoil holey carbon grid (R 1.2/1.3, 300 mesh, copper) previously glow-discharged (45 s, 25 mA). The grid was blotted for 4 s with a force of 0 using a Vitrobot Mark IV (FEI, ThermoFisher Scientific) with the environment chamber set to 100% humidity and 4°C. The sample was vitrified in liquid ethane cooled by liquid nitrogen.

First, 7161 movies were collected on a 200 kV Glacios (FEI, ThermoFisher Scientific) electron microscope with a Gatan K2 Summit camera using the following parameters: magnification 36,000, total electron dose of approximately 38 e<sup>-</sup>/Å<sup>2</sup>, pixel size 1.125 Å, 40 frames of 0.1 s. To overcome preferred specimen orientation, a tilted data set composed of 3038 additional movies was collected at 30° using a total dose of 38.9 e<sup>-</sup>/Å<sup>2</sup>, pixel size 1.125 Å, 44 frames of 0.1 s. Further collection details can be found in Supporting Information: Table 2.

Data processing was carried out using CryoSPARC v3.3.1.<sup>50,51</sup> After patch motion correction and patch CTF estimation, nontilted and 30° angle-tilted micrographs were manually curated, based on relative ice-thickness, motion, average defocus, and CTF fit resolution, leading to the selection of 6563 and 2283 micrographs, respectively.

Particles from both data sets were extracted, particles sorted by 2D classification and an ab initio reconstruction followed by nonuniform refinement resulted into an initial map (Supporting Information: Figure 10) displaying secondary structure features but also streaking artifacts in the direction of overrepresented particle orientation. To reduce this problem, a second processing attempt made only use the tilted data set. The initial map was used for template-based picking, and particles were sorted by performing two rounds of ab-initio reconstruction with three classes. The selected 221,316 particles were subjected to nonuniform refinement, followed by local motion correcting and another round of nonuniform refinement, resulting in a final map with an overall resolution of 4.5 Å. The monomer map was also generated using the tilted data set. Two rounds of ab initio reconstruction with three classes were performed resulting in the monomer map (from 86,635 particles).

## Structure prediction and docking

NDST1-4 monomer models predicted by AlphaFold2<sup>27,52</sup> were retrieved from AlphaFold Protein Structure Database (<https://alphafold.ebi.ac.uk/>, AlphaFold DB version 2022-11-01).

AlphaFold2<sup>28</sup> was used to generate NDST1 (residues 65–882) dimer. NDST1 dimer was placed in the electron density using ChimeraX.<sup>53</sup> After evaluating fit of the model into the density map in Coot,<sup>54</sup> residues 65–79 and 65–76 of subunit 1 and 2, respectively, were removed from the coordinates file. NDST1 dimer model was then refined with FlexEM<sup>29</sup> in CCP-EM<sup>55</sup> using the following parameters: ribfind cutoff 100, resolution 5.0 and five iterations.

UCFS Chimera version 1.16,<sup>56</sup> and ChimeraX version 1.2<sup>53</sup> were used to visualize and draw figures. NDST1 homodimer interface was analyzed with ChimeraX and PDBEPIA.<sup>57</sup>

## Surface plasmon resonance

Interactions between NDST1/NDST2 and LMW heparosan (CHE-BI:62819) were analyzed using surface plasmon resonance (SPR) on a

Biacore T200 system (Cytiva, ISBG) on a series S sensor chip NTA in HBS-P+ (Cytiva, 10 mM HEPES, 150 mM NaCl, pH 7.4, 0.05% P20). Following manufacturer's instructions, the surface was conditioned using a flow rate of 10  $\mu\text{L}/\text{min}$  by 60 s injection of 350 mM EDTA followed by 60 s 0.5 mM  $\text{NiCl}_2$  and 60 s of 3 mM EDTA. NDSTs were captured at 40 nM by their 8-His tag on the surface for 180 s reaching a capture level of 3400 and 1300 RU for NDST1 and NDST2, respectively. Single-cycle kinetics experiments were performed at 30  $\mu\text{L}/\text{min}$  using successive injections of increasing concentrations of LMW heparosan prepared in twofold serial dilutions, during 180 s (6.1–97  $\mu\text{M}$  for NDST1 and 64–1026  $\mu\text{M}$  for NDST2) in HBS-P+. Reference flow cells without  $\text{Ni}^{2+}$  activation nor ligand were used to subtract nonspecific binding. Affinities were measured using the steady-state analysis with the response level reached 60 s after the beginning of the injection with Biacore T200 Evaluation Software 3.2.1.

## Bio-Layer Interferometry

BLI experiments were performed on an Octet RED96e (Pall/ForteBIO/Sartorius, ISBG). HP (CHEBI:28304, H3393; Sigma-Aldrich) and HS (CHEBI:28815, HI-11098; Celsus Laboratories) were biotinylated on their reducing end with a PEG4 linker, as already described<sup>49</sup> and captured at 0.25  $\mu\text{g}/\text{mL}$  in Hepes buffer (10 mM HEPES, 150 mM NaCl, pH 7.4) for 600 s on SA (Streptavidin) biosensors previously hydrated in water.

Kinetics analysis were performed at 25°C in HBS-P+ (Cytiva), and a 600 s association step was followed by 600 s of free dissociation in HBS-P+. Six serial dilutions of each protein were used (1.56–50 and 3.13–100 nM for NDST1 and NDST2, respectively).

Finally, sensors were regenerated by washing with four or five cycles of 2 M NaCl for NDST1 and NDST2, respectively. Duplicates of the same interaction were performed on the same sensors.

Double referencing subtraction was applied using reference sensors on which no GAGs were loaded and reference wells that only contained HBS-P+. Experimental data were globally fitted to the 1:1 model with the Data Analysis HT 11 software (ForteBIO/Sartorius).

## AUTHOR CONTRIBUTIONS

Conceptualization, Hugues Lortat-Jacob and RW; Jerome Hénault, Christine Le Narvor, and David Bonnaffé prepared fluorescent oligosaccharide substrate; Emeline Richard and Bernard Priem prepared and analyzed LMW and HMW heparosans; protein expression and purification and enzymatic assays, Thibault Annaval and Sylvain D. Vallet; biophysical characterization and binding assays: Sylvain D. Vallet; cryo-EM data collection and processing, Rebekka Wild; data interpretation: Thibault Annaval, Sylvain D. Vallet, Rebekka Wild, Romain R. Vives, and Hugues Lortat-Jacob; supervision and funding acquisition, Rebekka Wild and Hugues Lortat-Jacob; Sylvain D. Vallet, Rebekka Wild, and Hugues Lortat-Jacob wrote the manuscript, all authors revised the manuscript.

## ACKNOWLEDGMENTS

The authors are grateful to Rida Awad, Yoan Monneau, and Damien Maurin for cloning and initial expression test of SEAP-NDST. The IBS Electron Microscope facility is supported by the Auvergne Rhône-Alpes Region, the Fonds Feder, the Fondation pour la Recherche Médicale and GIS-IBISA. We thank Caroline Mas for assistance and access to the Biophysical platform. Authors acknowledge Eleftherios Zarkadas for data collection and the other staff of the Electron Microscopy Platform from the Institut de Biologie Structurale and SPR/BLI platform scientific responsible, Jean-Baptiste Reiser, for their help and assistance. This work was granted access to the CCRT High-Performance Computing (HPC) facility under the Grant CCRT2022-login awarded by the Fundamental Research Division (DRF) of CEA. This work was supported by The Agence Nationale de la Recherche (Grant ANR-18-CE11-0006-01 to H. L. J. and Grant ANR-21-CE29-0022 to R. R. V.), the ATIP-Avenir program, and the “Investissements d'avenir” program Glyco@Alps, grant number [ANR-15-IDEX-02] to R. W. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003).

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

The NDST1 homodimer structure determined by cryoelectron microscopy was deposited in Electron Microscopy Data Bank (EMDB) under the accession code EMD-17349.

## ETHICS STATEMENT

The authors confirm that the ethical policies of the journal, as noted on the journal's author guidelines, have been adhered to.

## ORCID

Hugues Lortat-Jacob  <http://orcid.org/0000-0002-9772-0762>

## REFERENCES

- [1] S. D. Vallet, C. Berthollier, S. Ricard-Blum, *Am. J. Physiol. Cell Physiol.* **2022**, 322, C1271.
- [2] C. Marques, C. A. Reis, R. R. Vivès, A. Magalhães, *Front. Oncol.* **2021**, 11, 778752.
- [3] D. Shi, A. Sheng, L. Chi, *Front. Mol. Biosci.* **2021**, 8, 639666.
- [4] T. Annaval, R. Wild, Y. Créton, R. Sadir, R. R. Vivès, H. Lortat-Jacob, *Molecules* **2020**, 25, 4215.
- [5] Q. Wang, L. Chi, *Polymers* **2022**, 14, 5014.
- [6] R. A. Townley, H. E. Bülow, *Curr. Opin. Struct. Biol.* **2018**, 50, 144.
- [7] J. Kemna, E. Gout, L. Daniau, J. Lao, K. Weißert, S. Ammann, R. Kühn, M. Richter, C. Molenda, A. Sporbert, D. Zocholl, R. Klopffleisch, H. Lortat-Jacob, P. Aichele, T. Kammertoens, T. Blankenstein, *Nat. Immunol.* **2023**, 24, 414.
- [8] B. J. Connell, R. Sadir, F. Baleux, C. Laguri, J.-P. Kleman, L. Luo, F. Arenzana-Seisdedos, H. Lortat-Jacob, *Sci. Signal.* **2016**, 9, ra107.

- [9] D. Xu, K. Arnold, J. Liu, *Curr. Opin. Struct. Biol.* **2018**, *50*, 155.
- [10] F. E. Poulain, H. J. Yost, *Development* **2015**, *142*, 3456.
- [11] C. Laguri, N. Sapay, J.-P. Simorre, B. Brutscher, A. Imberty, P. Gans, H. Lortat-Jacob, *J. Am. Chem. Soc.* **2011**, *133*, 9642.
- [12] M. C. Z. Meneghetti, A. J. Hughes, T. R. Rudd, H. B. Nader, A. K. Powell, E. A. Yates, M. A. Lima, *J. R. Soc. Interface* **2015**, *12*, 20150589.
- [13] Y.-H. Chen, Y. Narimatsu, T. M. Clausen, C. Gomes, R. Karlsson, C. Steentoft, C. B. Sphlid, T. Gustavsson, A. Salanti, A. Persson, A. Malmström, D. Willén, U. Ellervik, E. P. Bennett, Y. Mao, H. Clausen, Z. Yang, *Nat. Methods* **2018**, *15*, 881.
- [14] C. McCormick, G. Duncan, K. T. Goutsos, F. Tufaro, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 668.
- [15] M. Busse-Wicher, K. B. Wicher, M. Kusche-Gullberg, *Matrix Biol.* **2014**, *35*, 25.
- [16] F. Leisico, J. Omeiri, C. Le Narvor, J. Beaudouin, M. Hons, D. Fenel, G. Schoehn, Y. Couté, D. Bonnaffé, R. Sadir, H. Lortat-Jacob, R. Wild, *Nat. Commun.* **2022**, *13*, 7110.
- [17] J.-P. Li, M. Kusche-Gullberg, *Int. Rev. Cell. Mol. Biol.* **2016**, *325*, 215.
- [18] J. Kreuger, L. Kjellén, *J. Histochem. Cytochem.* **2012**, *60*, 898.
- [19] P. Carlsson, J. Presto, D. Spillmann, U. Lindahl, L. Kjellén, *J. Biol. Chem.* **2008**, *283*, 20008.
- [20] J. Sheng, R. Liu, Y. Xu, J. Liu, *J. Biol. Chem.* **2011**, *286*, 19768.
- [21] Å. Hagner-McWhirter, U. Lindahl, J. Li, *Biochem. J.* **2000**, *347*(pt 1), 69.
- [22] D. S. Pikas, I. Eriksson, L. Kjellén, *Biochemistry* **2000**, *39*, 4552.
- [23] K. Grobe, J. Ledin, M. Ringvall, K. Holmborn, E. Forsberg, J. D. Esko, L. Kjellén, *Biochim. Biophys. Acta* **2002**, *1573*, 209.
- [24] A. Dagäl, A. Lundequist, B. Filipek-Górniok, T. Dierker, I. Eriksson, L. Kjellén, *Methods Mol. Biol.* **2022**, *2303*, 139.
- [25] J. van den Born, *Glycobiology* **2003**, *13*, 1.
- [26] D. F. Zielinska, F. Gnad, J. R. Wiśniewski, M. Mann, *Cell* **2010**, *141*, 897.
- [27] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Židek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583.
- [28] R. Evans, M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Židek, R. Bates, S. Blackwell, J. Yim, O. Ronneberger, S. Bodenstein, M. Zielinski, A. Bridgland, A. Potapenko, A. Cowie, K. Tunyasuvunakool, R. Jain, E. Clancy, P. Kohli, J. Jumper, D. Hassabis, *bioRxiv* **2022**. <https://doi.org/10.1101/2021.10.04.463034>
- [29] M. Topf, K. Lasker, B. Webb, H. Wolfson, W. Chiu, A. Sali, *Structure* **2008**, *16*, 295.
- [30] Y. Kakuta, T. Sueyoshi, M. Negishi, L. C. Pedersen, *J. Biol. Chem.* **1999**, *274*, 10673.
- [31] L. Holm, *Nucleic Acids Res.* **2022**, *50*, W210.
- [32] N. C. Bamford, F. Le Mauff, J. C. Van Loon, H. Ostapska, B. D. Snarr, Y. Zhang, E. N. Kitova, J. S. Klassen, J. D. C. Codée, D. C. Sheppard, P. L. Howell, *Nat. Commun.* **2020**, *11*, 2450.
- [33] J. Riesenfeld, M. Höök, U. Lindahl, *J. Biol. Chem.* **1980**, *255*, 922.
- [34] Z. Wei, S. J. Swiedler, *J. Biol. Chem.* **1999**, *274*, 1966.
- [35] G. Fan, L. Xiao, L. Cheng, X. Wang, B. Sun, G. Hu, *FEBS Lett.* **2000**, *467*, 7.
- [36] E. Forsberg, G. Pejler, M. Ringvall, C. Lunderius, B. Tomasini-Johansson, M. Kusche-Gullberg, I. Eriksson, J. Ledin, L. Hellman, L. Kjellén, *Nature* **1999**, *400*, 773.
- [37] D. E. Humphries, G. W. Wong, D. S. Friend, M. F. Gurish, W. T. Qiu, C. Huang, A. H. Sharpe, R. L. Stevens, *Nature* **1999**, *400*, 769.
- [38] L. Kjellén, *Biochem. Soc. Trans.* **2003**, *31*, 340.
- [39] A. Deligny, T. Dierker, A. Dagäl, A. Lundequist, I. Eriksson, A. V. Nairn, K. W. Moremen, C. L. R. Merry, L. Kjellén, *J. Biol. Chem.* **2016**, *291*, 18600.
- [40] A. F. Moon, Y. Xu, S. M. Woody, J. M. Krahn, R. J. Linhardt, J. Liu, L. C. Pedersen, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5265.
- [41] W. Dou, Y. Xu, V. Pagadala, L. C. Pedersen, J. Liu, *J. Biol. Chem.* **2015**, *290*, 20427.
- [42] P. Missaghian, T. Dierker, E. Khosrowabadi, F. Axling, I. Eriksson, A. Ghanem, M. Kusche-Gullberg, S. Kellokumpu, L. Kjellén, *Glycobiology* **2022**, *32*, 518.
- [43] R. Liu, Y. Xu, M. Chen, M. Weiwer, X. Zhou, A. S. Bridges, P. L. DeAngelis, Q. Zhang, R. J. Linhardt, J. Liu, *J. Biol. Chem.* **2010**, *285*, 34240.
- [44] B. Priem, J. Peroux, P. Colin-Morel, S. Drouillard, S. Fort, *Carbohydr. Polym.* **2017**, *167*, 123.
- [45] H. Barreteau, E. Richard, S. Drouillard, E. Samain, B. Priem, *Carbohydr. Res.* **2012**, *360*, 19.
- [46] Z. Wang, B. Yang, Z. Zhang, M. Ly, M. Takiuddin, S. Mousa, J. Liu, J. S. Dordick, R. J. Linhardt, *Appl. Microbiol. Biotechnol.* **2011**, *91*, 91.
- [47] K. Raman, S. Arungundram, *Glycosaminoglycans: Methods and Protocols* (Eds: K. Balagurunathan, H. Nakato, U. Desai, Y. Saijoh), Springer, New York **2022**, p. 289.
- [48] E. Smeds, A. Feta, M. Kusche-Gullberg, *Glycobiology* **2010**, *20*, 1274.
- [49] C. Laguri, R. Sadir, E. Gout, R. R. Vivès, H. Lortat-Jacob, *Methods Mol. Biol.* **2022**, *2303*, 121.
- [50] A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, *Nat. Methods* **2017**, *14*, 290.
- [51] A. Punjani, H. Zhang, D. J. Fleet, *Nat. Methods* **2020**, *17*, 1214.
- [52] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Židek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, *Nucleic Acids Res.* **2022**, *50*, D439.
- [53] E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, T. E. Ferrin, *Protein Sci.* **2021**, *30*, 70.
- [54] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486.
- [55] T. Burnley, C. M. Palmer, M. Winn, *Acta Crystallogr. D Struct. Biol.* **2017**, *73*, 469.
- [56] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605.
- [57] E. Krissinel, K. Henrick, *J. Mol. Biol.* **2007**, *372*, 774.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** S. D. Vallet, T. Annaval, R. R. Vives, E. Richard, J. Hénault, C. Le Narvor, D. Bonnaffé, B. Priem, R. Wild, H. Lortat-Jacob, *Proteoglycan Res.* **2023**, e8.  
<https://doi.org/10.1002/pgr2.8>