Acta Cryst. (1998). D**54**, 111–113

Crystallization and preliminary X-ray studies of sialidase L from the leech *Macrobdella decora*

YU LUO,[a] MIN-YUAN CHOU,[b] SU-CHEN LI,[b] YU-TEH LI[b] and MING LUO[c]* *at*
[a]*Center for Macromolecular Crystallography, University of Alabama at Birmingham, Alabama 35294, USA,* [b]*Department of Biochemistry, Tulane University School of Medicine, New Orleans, Louisiana 70112, USA, and* [c]*Center for Macromolecular Crystallography, Department of Microbiology, University of Alabama at Birmingham, Alabama 35294, USA.*
E-mail: ming@orion.cmc.uab.edu

(*Received 11 February 1997; accepted 27 May 1997*)

Abstract

Functional monomeric 83 kDa sialidase L, a NeuAcα2→3Gal-specific sialidase from *Macrobdella* leech, was expressed in *Escherichia coli* and readily crystallized by a macroseeding technique. The crystal belongs to space group *P*1 with unit-cell parameters a = 46.4, b = 69.3, c = 72.5 Å, α = 113.5, β = 95.4 and γ = 107.3°. There is one molecule per unit cell, giving a V_m = 2.4 Å^3 Da^{-1} and a solvent content of 40%. Native and mercury-derivative data sets were collected to 2.0 Å resolution. Threading and molecular-replacement calculations confirmed the existence of a bacterial sialidase-like domain.

1. Introduction

Sialidase (E.C. 3.2.1.18) hydrolyzes α-ketosidically linked sialic acids from sialoglycoconjugates. They are widely distributed in nature and have been isolated from microorganisms as well as mammalian tissues (Schauer & Vliegenthart, 1982). Bacterial sialidases have been suggested to play a role in initiating infections of animals (Saito & Yu, 1995; Schauer, 1983), while mammalian sialidases have been shown to be involved in the sialoglycoconjugate catabolism (Schauer & Vliegenthart, 1982).

Sialidase L, an 83 kDa sialidase found in *Macrobdella* leech, is the first sialidase found to show a strict specificity towards the hydrolysis of NeuAcα2→3Gal linkage that releases 2,7-anhydro-NeuAc instead of NeuAc from sialoglycoconjugates. Therefore, it may also become useful for selective cleavage of sialoglycoconjugates without destroying other sialosyl linkages, such as NeuAcα2→6Gal, NeuAcα2→6GalNAc, NeuAcα2→6GlcNAc, NeuAcα2→8NeuAc and NeuAcα2→9NeuAc. Interestingly, 2-deoxy-2,3-dehydro-NeuAc, the potent competitive inhibitor of microbial and mammalian sialidases, has little inhibitory effects on the activity of sialidase L (Chou, Li, Kiso, Hasegawa & Li, 1994). The amino-acid sequence of sialidase L shows an 'FRIP' motif and four 'Asp boxes', Ser-*X*-Asp-*X*-Gly-*X*-Thr-Trp, which are conserved motifs in reported bacterial and mammalian sialidases (Roggentin, Schauer, Hoyer & Vimr 1993; Roggentin *et al.*, 1989). Further sequence analysis showed sialidase L may have an N-terminal domain and a catalytic domain which resembles the bacterial sialidases (Chou *et al.*, 1994; Chou, Li & Li, 1996). We expect that the high-resolution crystal structure will eventually elucidate the catalytic mechanism of sialidase L as well as its unique specificity.

2. Experimental

Functional sialidase L was expressed in *Escherichia coli* and purified by the procedure described by Chou *et al.* (1996). The open reading frame encodes 762 amino acids with a calculated molecular mass of 82 982 Da.

The initial screening in ammonium sulfate, sodium sulfate, sodium citrate, or PEG 2K to PEG 20K gave ~0.05 mm crystals as a dense precipitate. Some crystals produced diffraction patterns to 4 Å resolution. To improve crystal growth, samples were mixed with an equal volume of 50 m*M* 2-mercaptoethanol and stored at 277 K for two weeks. Gel filtration with a Superdex-200 (Pharmarcia) column using 0.15 *M* (pH = 7.0) sodium phosphate buffer as the eluent, gave satisfactory separation between the ~80 kDa monomer fraction and the ~160 kDa fraction, which is also active. The fact that the sample untreated with 2-mercaptoethanol predominantly yielded a dimer, and that the polypeptide chain has five cysteines, suggests that the ~160 kDa fraction may be a dimer with possible intermolecular disulfide bonds formed by free cysteines. The monomer fraction was repeatedly concentrated with a CENTRICON-30 at 4500*g* and diluted with 20 m*M* 2-mercaptoethanol to remove the counter ions. It was then subjected to a finer screening with PEG 6K at both 295 and 277 K. The best seeds were harvested in a month from hanging drops set up at 277 K with 2 µl of ~15 mg ml^{-1} protein and 2 µl 20% PEG 6K in 0.1 *M* cacodylate buffer containing 0.25 *M* NaCl at pH 6.3 over the same precipitant solution. A macroseeding technique was employed by adding 2–3 seeds with 2 µl of precipitant solution mentioned above to 4 µl of ~2 mg ml^{-1} protein solution. Crystals grow to a typical size of 0.3 × 0.2 × 0.1 mm in 1–2 weeks (Fig. 1). The crystals diffract typically to at least 2.0 Å resolution and were used to collect X-ray diffraction data. The crystals are triclinic in space group *P*1, with unit-cell parameters a = 46.4, b = 69.3, c = 72.5 Å, α = 113.5, β = 95.4 and γ = 107.3°. Assuming there is one molecule per unit cell, a value for V_m = 2.4 Å^3 Da^{-1} and a solvent content of 40% were obtained, which are within the normal range for protein crystals (Matthews, 1968). Another crystal form was readily grown by the same procedure at 295 K with 15% PEG 6K in 0.2 *M* pH = 5.2 acetate buffer. The second crystalline form, also in space group *P*1, but with approximately a three times larger cell volume (a = 72.6, b = 73.3, c = 128.0 Å, α = 78.8, β = 77.0 and γ = 61.8°), has been used only in initial testing of the reactivity of heavy-atom reagents because of their much larger mosaicity and poorer diffraction to ~3.0 Å resolution.

The X-ray diffraction analysis was performed on a DIP-2030 image-plate detector system (Mac Science) with focusing mirror optics mounted on an RU-200 rotating-anode X-ray generator (Rigaku) operating at 50 kV and 100 mA. Data were collected at 100 K in a nitrogen gas stream (Oxford Cryosystems Cryostream) using cryoprotectant made with 2 vol glycerol and 8 vol precipitant. The detector was set at

150 mm from the crystal, and 240 frames of 1.2° oscillation images with a 1800 s exposure time were acquired using one crystal each for both the native and a mercuric derivative. The derivative crystal was prepared by soaking the crystal in 0.4 m*M* mercuric acetate for 3 h. After 120 frames, the swing angle of the goniometer was set to 10° to get wider coverage of the data. No significant decay was observed during ~5 d of data collection. The data sets were processed with the *HKL* package (Otwinowski, 1993; Minor, 1993). The native data are 96.5% complete to 2.0 Å with $R_{\rm merge}$ = 4.4% (Table 1). The corresponding values for the 2.0 Å mercury-derivative data are 92.5 and 6.2%. The $R_{\rm iso}$ is 16.5%. No useful anomalous scattering information was observed.

Table 1. *Data-collection statistics*

	Native	$Hg(Ac)_2$
Resolution (Å)	2.0	2.0
Observations	145108	125513
Unique reflections	48481	47005
Completeness	96.5	92.5
$\langle I\rangle/\langle\sigma I\rangle$	16.0	13.8
$R_{\rm merge}$ (%)†	4.4	6.2
$R_{\rm iso}$ (%)‡	—	16.5

† $R_{merge} = \sum_{hkl}\sum_i(|I_{hkl} - \langle I_{hkl}\rangle|)/\sum_{hkl,i}\langle I_{hkl}\rangle$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with indices *hkl* and $\langle I_{hkl}\rangle$ is the mean intensity of that reflection. ‡ The mean isomorphous difference $R_{iso} = \sum_{hkl}(||F_{hkl,\rm P}| - |F_{hkl,\rm PH}||)/\sum|F_{hkl,\rm P}|$ where $F_{hkl,\rm P}$ is the structure factor of a reflection from the native data set and $F_{hkl,\rm P}$ is the structure factor of the corresponding reflection from the native data set.

3. Results and discussion

Since a catalytic domain homologous to bacterial sialidase has been proposed for sialidase L (Chou *et al.*, 1994, 1996), *THREADER* (Jones, Taylor & Thornton, 1992) was used to identify the most similar fold amongst the *THREADER*'s fold library with several β-propeller structures added. The top hit was a 41 kDa sialidase from *Micromonospora viridifaciens* (Gaskell, Crennell & Taylor, 1995) with a *Z* score = −4.28 against 353 folds. Its sequence identity with sialidase L is 18%. Other folds that top the scoring list are either sialidases or G-protein β domains which also have a β-propeller architecture (Wall *et al.*, 1995; Lambright *et al.*, 1996; Sondek, Bohm, Lambright, Hamm & Sigler, 1996).

Molecular replacement was attempted with *AMoRe* (Navaza, 1994) of the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). A mediocre solution with a correlation coefficient of 25.9%, against a continuum of noise value starting from 20%, was found using the 361-amino-acid *M. viridifaciens* sialidase structure as a full atom search model. Search models constructed from a few other related proteins gave best correlation coefficients of no better than 22% against a similar noise spectrum. The fitted model, positioned in the sialidase L unit cell, was refined with 6.0–2.0 Å resolution data to an *R* value of 0.44 using *X-PLOR* (Brünger, 1992). The phases obtained from the refined partial model were then used to locate the heavy atoms by difference-Fourier synthesis. Two Hg sites (0.181, 0.688, 0.977 and 0.718, 0.209, 0.677, 11σ and 8σ above the average density level, respectively) were located in the 15.0–2.0 Å difference Fourier map. The interatomic vector of the two sites is consistent with the only distinguishable non-origin peak at (0.463, 0.482, 0.302, 10σ) in the 15.9–2.0 Å isomorphous-difference Patterson map (Fig. 2). The occupancies of the two sites were refined with *MLPHARE/CCP*4 to 1.11 and 0.28, respectively. It is likely that a change in the soaking time may give a second useful derivative. Combination of SIR phases and MR phases has been unsuccessful in obtaining electron densities for the other ~400 amino-acid residues.

The existence of a bacterial sialidase-like catalytic domain indicated by sequence patterns has been supported by threading and the molecular replacement solution using the sialidase from *M. viridifaciens* as the search model. This has been confirmed by the successful location of two heavy-atom sites using

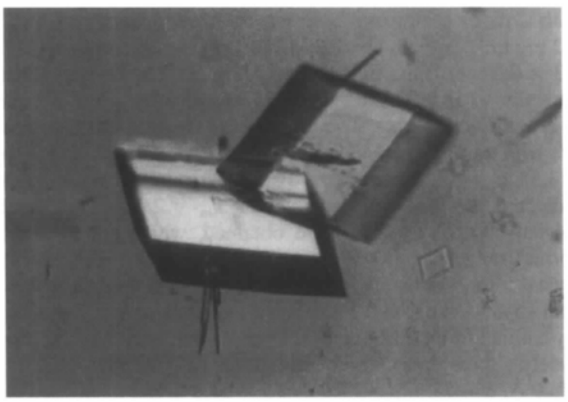

Fig. 1. A photograph of the sialidase L triclinic crystals (~75×) in the hanging drop. The crystals were obtained by a macroseeding technique: 2–3 ~0.05 mm microcrystals were introduced with 2 µl precipitant solution to 4 µl of ~2 mg ml^{-1} sialidase L solution. The hanging drops were set up on ice and then transfered into a 277 K incubator. The crystals grew to maximum sizes of about 0.3 × 0.2 × 0.1 mm during a period of 1–2 weeks. The optimum precipitant contains 20% PEG 6K in 0.1 *M* cacodylate buffer and 0.25 *M* NaCl. The ~0.05 mm seeds originally came from autonucleation. Later on, more reproducible seeds were harvested by seeding with 1000× diluted solution containing crunched crystals.

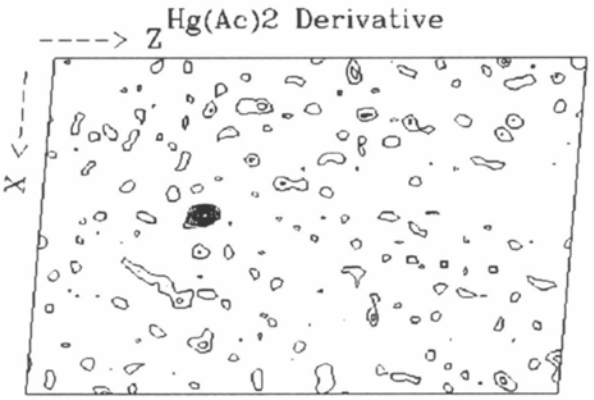

Fig. 2. The isomorphous-difference Patterson map of the $Hg(Ac)_2$ derivative using 15.0–2.0 Å resolution data. The map was contoured at Y = 0.48 section with 1σ interval starting from the average density value of the map. Only one non-origin peak, together with its inversed copy, was identified, ~10σ above the average density level, at X = 0.463, Y = 0.482, Z = 0.302. The second highest non-origin peak was only ~5σ above average. The identified peak position is consistent with the two Hg sites (0.181, 0.688, 0.977 and 0.718, 0.209, 0.677, respectively) found by difference Fourier synthesis with phases from a molecular-replacement solution.

molecular-replacement phases. Screening of potential heavy-atom derivatives is still in progress. The crystals of sialidase L with high-resolution diffracting quality open the possibility of solving its structure at the atomic resolution level and elucidating its unique specificity towards NeuAcα2$\rightarrow$3Gal linkage and its enzymatic mechanism on a structural basis.

We are grateful to Dr Susan J. Crennell who provided us with the coordinates of bacterial sialidases, and to Professor Stephan R. Sprang and Dr John Sondek for their latest G-protein coordinates. We would also thank Dr Bingdong Sha for helpful discussions.

References

Brünger, A. T. (1992). *X-PLOR Manual* 3.1, Yale University, New Haven, CT, USA.

Chou, M. Y., Li, S. C., Kiso, M., Hasegawa, A. & Li, Y. T. (1994). *J. Biol. Chem.* **269**, 18821–18826.

Chou, M. Y., Li, S. C. & Li, Y. T. (1996). *J. Biol. Chem.* **271**, 19219–19224.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Gaskell, A., Crennell, S. & Taylor, G. (1995). *Structure*, **3**, 1197–1205.

Jones, D. T., Taylor, W. R. & Thornton, J. M. (1992). *Nature (London)*, **358**, 86–89.

Lambright, D. G., Sondek, J., Bohm, A., Skiba, N. P., Hamm, H. E. & Sigler, P. B. (1996). *Nature (London)*, **379**, 311–319.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Minor, W. (1993). *XdisplayF Program*. Purdue University, West Lafayette, USA.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*, pp. 56–62. Warrington: Daresbury Laboratory.

Roggentin, P., Rothe, B., Kaper, J. B., Galen, J., Lawrisuk, L., Vimr, E. R. & Schauer, R. (1989). *Glycoconjugate J.* **6**, 349–353.

Roggentin, P., Schauer, R., Hoyer, L. L. & Vimr, E. R. (1993). *Mol. Microbiol.* **9**, 915–921.

Saito, M. & Yu, R. K. (1995). *Biology of the Sialic Acids: Biochemistry and Function of Sialidases*, edited by A. Rosenberg, pp. 261–313. New York: Plenum Press.

Schauer, R. (1983). *Structural Carbohydrates of the Liver: Glycosidases with Special Reference to the Pathophysiological Role of Sialidases*, edited by H. Popper, W. Reutter, F. Gudat & E. Kotten, pp. 83–97. Baltimore: University Park Press.

Schauer, R. & Vliegenthart, J. F. G. (1982). *Sialic Acids: Chemistry, Metabolism and Function*, edited by R. Schauer, pp. 1–3. Vienna: Springer-Verlag.

Sondek, J., Bohm, A., Lambright, D. G., Hamm, H. E. & Sigler, P. B. (1996). *Nature (London)*, **379**, 369–374.

Wall, M. A., Coleman, D. E., Lee, E., Iniguez-Lluhi, J. A., Posner, B. A., Gilman, A. G. & Sprang, R. (1995). *Cell*, **83**, 1047–1058.