Cloning, expression, purification, crystallization and preliminary X-ray analysis of NodS N-methyltransferase from *Bradyrhizobium japonicum* WM9

Ozgur Cakici[a], Michal Sikorski[a], Tomasz Stepkowski[a] , Grzegorz Bujacz[ab] and Mariusz Jaskolski[ac]*

[a]*Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland,* [b]*Institute of Technical Biochemistry, Faculty of Biotechnology and Food Sciences, Technical University of Lodz, Lodz, Poland and* [c]*Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, Poland. E-mail: mariuszj@amu.edu.pl*

Abstract

The Nod factor (NF) is a rhizobial signal molecule involved in the recognition of a legume host and formation of root and stem nodules. There are some unique enzymes involved in the biosynthesis of NF, which is a variously but specifically substituted lipo-chito-oligosaccharide. One of the enzymes is NodS, an N-methyltransferase that methylates end-deacetylated chito-oligosaccharide substrates. For the methylation reaction, NodS uses S-adenosyl-L-methionine (SAM) as methyl donor. So far, there is no structural information about NodS from any rhizobium. We have undertaken X-ray crystallographic studies of the NodS protein from *Bradyrhizobium japonicum* WM9, which infects lupin and serradella legumes. The *nodS* gene was cloned and the recombinant protein was expressed in *E. coli* cells, both using natural amino acids and as a Se-Met derivative. NodS without ligands was crystallized in the presence of PEG3350 and $MgCl_2$. The protein was also crystallized in complex with S-adenosyl-L-homocysteine (SAH) in the presence of PEG8000 and $MgCl_2$. SAH is a byproduct of the methylation reaction, produced from SAM. The crystals of apo NodS are tetragonal and diffract X-rays to 2.42 Å resolution. The NodS-SAH complex crystallizes in an orthorhombic space group and the crystals diffract X-rays to 1.85 Å.

Keywords: *Bradyrhizobium*, Nod Factor, NodS, N-methyltransferase, SAM, SAH, symbiosis, legume plants, nitrogen assimilation

1 Introduction

An exchange of precisely tuned low-molecular-weight signal molecules between rhizobia and legume plants results in the establishment of a highly specific symbiotic association. During this process, the

rhizobia colonize the roots, forming specialized plant organs, the root nodules, in which they assimilate atmospheric nitrogen. Specific flavonoid signals secreted by the legume roots induce the transcription of some rhizobial genes (Long, 2001). Most of these genes, referred to as the *nod*, *nol* and *noe* genes, are responsible for the synthesis and release of the bacterial signal molecules, known as Nod (nodulation) Factors (NF), whose recognition by the plant host leads to induction of the nodule meristem and consequently to nodule formation. At the same time, the rhizobium cells enter the root tissues either through root hairs or cracks in the root epidermis, colonizing the newly formed nodule cells, in which they undergo transformation into nitrogen fixing bacteroids. The Nod factors contain a common lipo-chito-oligosaccharide core structure that carries a number of substitutions, specific for different rhizobial strains. This structural diversity of the Nod factors is an essential determinant of host specificity. There are numerous enzymes involved in the assembly and modification of the Nod factor during its biosynthesis (Denarie *et al.*, 1992; Carlson *et al.*, 1994), some of which are only found in rhizobia. One of them is NodS, which functions as N-methyltransferase, modifying by SAM (S-adenosyl-L-methionine)-dependent methylation the non-reducing end of deacetylated chito-oligosaccharide substrates (Geelen *et al.*, 1993; Geelen *et al.*, 1995). In common with all SAM-dependent methylation reactions, NodS converts SAM to SAH (S-adenosyl-L-homocysteine), which is a byproduct of the catalysis (Fig. 1). NodS is the first example of an SAM-dependent methyltransferase that methylates a chito-oligosaccharide substrate.

The currently available information about the three-dimensional structure of the enzymes involved in NF synthesis is extremely limited (Brzezinski *et al.*, 2007). In particular, there are no structural data about NodS from any rhizobium strain. We have, therefore, undertaken X-ray crystallographic studies of the 210-residue NodS protein from *Bradyrhizobium japonicum* WM9. Strain WM9 infects lupins and serradella, and its nodulation gene sequences belong to a distinct group classified as Clade II (Stepkowski *et al*. 2003; 2007). This project is part of our wider program aimed at structural characterization of all the enzymes involved in the Nod factor biosynthetic pathway.

2 Materials and Methods

2.1 Cloning

The *nodS* gene (AF222753) was amplified from genomic DNA of *Bradyrhizobium japonicum* WM9 strain *via* polymerase chain reaction (PCR). The primers used were: forward 5'-CACCATGGTGAGCGTAGACAACAC-3' and reverse 5'-TCAAGCTCGTCCGTCGGGG-3'. The forward primer contained four additional nucleotides (CACC) at the 5' end, recognized by the TOPO polymerase that is covalently attached to the ends of the pET151-TOPO expression vector (Invitrogen), to generate a sticky overhang for subcloning of the PCR product in the correct orientation. Because of the

requirement for blunt-ended PCR products, amplification of the target gene sequence for cloning was performed using *Pwo* DNA Polymerase (Roche) with proofreading activity. The amplification reaction mixture (50 µl) consisted of 1x *Pwo* DNA polymerase buffer, 1.5 mM $MgSO_4$, 0.2 mM dNTPs, 0.02 U/µl *Pwo* DNA Polymerase, 0.4 µM of each primer, and 80 ng of template DNA. The initial denaturation step was performed at 367 K for 5 minutes and samples were then subjected to 30 cycles of denaturation at 367 K for 1 minute, annealing at 328 K for 1 minute and extension at 345 K for 1 minute. Subsequently, the PCR was completed by a final extension step at 345 K for 10 minutes. Samples were stored at 277 K. The PCR products were analyzed by electrophoresis in 1% (w/v) agarose gels and detected by staining with ethidium bromide.

The amplified fragment was cloned into pET151/D-TOPO expression vector using Champion™ pET151 Directional TOPO Expression Kit (Invitrogen). The expression vector has an N-terminal His_6 tag, which is followed by a TEV protease cleavage site. The expression clone was confirmed by DNA sequencing and the vector was transformed into BL21-CodonPlus (DE3)-RIPL competent *E. coli* cells (Stratagene).

2.2 Expression and Purification

40 ml of TB medium containing 100 $mg \cdot l^{-1}$ ampicillin and 34 $mg \cdot l^{-1}$ chloramphenicol were inoculated with the transformant. The cells were grown at 310 K for 5-8 h till turbidity. The activated culture was transferred into 2 l of TB medium containing 100 $mg \cdot l^{-1}$ ampicillin and 34 $mg \cdot l^{-1}$ chloramphenicol. The cells were grown until the OD_{600} reached 0.8-1.2. The culture was cooled to 291 K and induced by isopropyl-β-D-thiogalactopyranoside (IPTG) at a final concentration of 0.2 mM. After induction, the culture was incubated overnight at 291 K. The cell paste was harvested and frozen on dry ice for storage at 193 K.

The same transformant was also used for the expression of Se-Met derivative protein. Se-Met medium base from Molecular Dimensions, Ltd. (MDL) and nutrient mix (MDL) were used as growth media. The cultivation procedures were as for the wild-type protein, except that the volumes were reduced by 50%. When the OD_{600} reached 0.8-1.2, the culture was cooled to 291 K and 10 ml of Met-Stop solution (100 $mg \cdot l^{-1}$ of Lys, Thr, Phe; 50 $mg \cdot l^{-1}$ of Ile, Val, Leu) were added to block the methionine biosynthesis pathway. The culture was supplemented with 4 ml of selenomethionine solution from MDL and after 15 min incubation at 291 K, induced by IPTG.

About 10 g of wet cell paste were resuspended in 50 ml of lysis buffer (50 mM Tris·HCl pH 7.5, 500 mM NaCl, 10% v/v glycerol, 2.5 mM TCEP, 1 mM PMSF, 1 $mg \cdot ml^{-1}$ lysozyme). After incubation on ice for 30-60 min, the lysate was sonicated for 90 sec on ice. Cell debris was removed by centrifugation. The supernatant was subjected to purification using an ÄKTA Purifier system (GE Healthcare). In the first step, the proteins were applied on a 1 ml HisTrap™ column. After binding, the column was washed with 30 mM imidazole in buffer A (50 mM Tris·HCl pH 7.5, 500 mM NaCl, 10% v/v glycerol, 2.5 mM TCEP).

The protein was eluted from the column using 20-25 ml of a linear 30-300 mM gradient of imidazole in buffer A. The eluted protein was desalted (HiPrep™ 26/10 column) using buffer A to remove imidazole. Imidazole-free protein solution was incubated overnight at 277 K with a His-tagged TEV protease (60 µg TEV protease / 1 mg His-tagged protein) to cleave off the His-tag. Subsequently, the sample was applied on a HisTrap™ column to remove the TEV protease, the His-tag and any undigested fusion protein. The first flow-through was collected and applied on a gel filtration column (HiLoad 16/60 Superdex 200 pg) at a 1 ml·min^{-1} flow rate in a buffer containing 50 mM Tris·HCl pH 7.5, 150 mM NaCl, 10% v/v glycerol, 5 mM DTT, and 1 mM EDTA. Peak fractions were collected and analyzed by SDS-PAGE and MALDI-TOF mass spectrometry (calculated M_r 24123 Da, observed 24120 Da). The purified protein was concentrated to 4 mg·ml^{-1} using Amicon Ultra-4 centrifugal devices with 10 kDa cut-off (Millipore) and stored in small aliquots at 193 K. All purification and concentration steps were carried out at 277 K. Protein concentration was determined spectrophotometrically at 280 nm or by the Bradford method (Bradford, 1976) with BSA as a standard. The purified protein has native *Bradyrhizobium japonicum* WM9 sequence (Stepkowski *et al*., 2003) with an additional hexapeptide at the N-terminus (GIDPFT) introduced as a cloning artifact.

2.3 Crystallization

Prior to setting up the crystallization screens, the protein sample was concentrated to 10 mg·ml^{-1} and the buffer was changed to 10 mM Tris·HCl pH 8.0, 100 mM NaCl, 2% v/v glycerol, 2.5 mM TCEP, using Ultrafree Centrifugal Filter Units (Millipore) with 10 kDa cut-off at 277 K, and the protein solution was passed through an Ultrafree-MC Centrifugal Filter Unit (Millipore) with 0.1 µm pore size at 277 K. The sitting-drop vapor-diffusion screening for initial crystallization conditions was performed for wild-type and Se-Met apo NodS in the high-throughput (HT) crystallization service unit at the EMBL, Hamburg, Germany (Mueller-Dieckmann, 2006). Crystal and Index Screens (Hampton Research) were used for the initial experiments. 200 nl protein samples were mixed with an equal amount of the reservoir solution and equilibrated against 50 µl reservoir solution, and the crystallization plates were stored at 292K. In the next step, optimum pH, precipitant, protein and additive concentration, and the drop size were adjusted in in-house experiments according to the promising results (one hit observed for wild-type apo NodS) of the HT screening. In the optimization screens, the protein sample was mixed with reservoir solution in 1:1 ratio and equilibrated in sitting- or hanging-drops against 1 ml reservoir solution at 292 K. The crystals appeared after one day and grew to their final dimensions in five days (Figure 2A). For co-crystallization with S-adenosyl-L-homocysteine, a protein sample at 4 mg·ml^{-1} was mixed with a stoichiometric amount of 2 mM SAH and incubated overnight at 277 K. Prior to setting up the crystallization screen, the sample was concentrated to 10 mg·ml^{-1} and filtrated using an Ultrafree-MC Centrifugal Filter Unit with 0.1 µm

pore size at 277 K. The crystals appeared after one day and grew to their final dimensions in seven days at 292 K (Figure 2B).

2.4 Data Collection and Processing

Synchrotron radiation was used to collect diffraction data for all crystals, after their flash-vitrification at 100 K in a stream of nitrogen gas. A ligand-free crystal of NodS was cryoprotected in a 1:1 mixture of the reservoir solution and 50% (v/v) PEG400. Diffraction data were collected at the EMBL beamline X13 of the DESY synchrotron in Hamburg, using a 165 mm MAR CCD detector (Figure 3A). Three-hundred 0.75° oscillation images were recorded at a 210 mm crystal-to-detector distance, corresponding to 2.3 Å resolution (Table 1). A crystal of native NodS in complex with SAH was cryoprotected using the reservoir solution supplemented with 14% (v/v) PEG400. Diffraction data were collected at beamline BL 14.2 of the BESSY synchrotron in Berlin using a 165 mm MAR CCD detector (Figure 3B). The diffraction data were collected in two runs, corresponding to low (2.3 Å) and high (1.71 Å) resolution, as specified in Table 1. An Se-Met NodS crystal grown in the presence of SAH was used for a three-wavelength MAD data collection at beamline BL 14.2 of the BESSY synchrotron in Berlin, equipped with a 165 mm MAR CCD detector. The crystal was cryoprotected in the reservoir solution supplemented with 12% (v/v) PEG400. A fluorescence scan at the Se *K*α absorption edge taken with a Roentec X-Flash detector was used to define the MAD wavelengths λ_{edge}= 0.97984, λ_{peak}= 0.97968 and λ_{remote}= 0.95000 Å (Table 1). Indexing, integration and scaling of all diffraction images were performed in *HKL2000* (Otwinowski & Minor, 1997).

3 Results and Discussion

Soluble recombinant NodS protein with *Bradyrhizobium japonicum* WM9 sequence extended with a short N-terminal tag introduced as a cloning artifact was obtained by overnight cultivation of transformed *E. coli* cells at 291 K. The protein was successfully purified by FPLC chromatography and its homogeneity and monomeric form were confirmed by gel filtration in the final purification step. The protein was also produced and purified in selenomethionyl form using non-auxotrophic *E. coli* cells and a cultivation protocol that blocks the methionine biosynthetic pathway. Successful Met→Se-Met substitution of the six Met sites was confirmed by MALDI-TOF mass spectrometry (calculated M_r 24405 Da, observed 24385 Da). Initial microcrystals of ligand-free NodS were found after HT crystallization screening. After optimization trials, the best diffracting crystals of ligand-free NodS were obtained at 292 K in 28% PEG3350, 0.1 M $MgCl_2$, pH 7.9. The best crystals of NodS in complex with SAH were obtained at 292 K in 16% PEG8000, 5 mM $MgCl_2$, pH 8.5, using the streak seeding method and protein concentration of 10

mg·ml^{-1}. The SAH complex of selenomethionyl NodS was crystallized in similar condition (4 mg·ml^{-1} protein concentration, in 14% PEG8000, 40 mM $MgCl_2$, pH 8.5, at 292K), also with the aid of streak seeding. The crystals of ligand-free NodS are tetragonal, space group $P4_122$ or $P4_322$, with one protein molecule in the asymmetric unit estimated at a Matthews coefficient (Matthews, 1968) of 1.74 Å^3·Da^{-1} (solvent content 29.2%). The NodS-SAH complex crystallizes in the orthorhombic space group $P2_12_12$ with four molecules in the asymmetric unit estimated at a Matthews coefficient of 2.28 Å^3·Da^{-1} (solvent content 46.6%). Structure determination is in progress.

Acknowledgement

This work was supported in part by a Marie Curie Early Stage Researcher Training (MEST-CT-2004-504066) fellowship of the 6th Framework Programme (FP6) of the European Union. The skillful help of Dr. Robert Kolodziejczyk and Dr. Krzysztof Brzezinski during the diffraction experiments is gratefully acknowledged.

References

Bradford, M.M. (1976). *Anal. Biochem.* **72**, 248-254.

Brzezinski, K., Stepkowski, T., Panjikar, S., Bujacz, G. & Jaskolski, M. (2007). *Acta Biochim. Polon.* **54**, 537-549.

Carlson, R.W., Price, N.P.J. & Stacey, G. (1994). *Mol. Plant Microbe Interact.* **7**, 684-695.

Denarie, J., Debelle, F. & Rosenberg, C. (1992). *Annu. Rev. Microbiol.* **46**, 497-531.

Evans, P. (2006). *Acta Cryst.* D**62**, 72-82.

Geelen, D., Mergaert, P., Geremia, R.A., Goormachtig, S., Van Montagu, M. & Holsters, M. (1993). *Mol Microbiol.* **9**, 145-154.

Geelen, D., Leyman, B., Mergaert, P., Klarskov, K., Van Montagu, M., Geremia, R. & Holsters, M. (1995). *Mol Microbiol.* **17**, 387-397.

Long, S.R. (2001). *Plant Physiol*. **125**, 69-72.

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

Mueller-Dieckmann, J. (2006). *Acta Cryst.* D**62**, 1446-1452.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307-326.

Stepkowski, T., Hughes, C.E., Law, I.J., Markiewicz, L., Gurda, D., Chlebicka, A. & Moulin, L. (2007). *Appl. Environ. Microbiol.* **73**, 3254-3264.

Stepkowski, T., Swiderska, A., Miedzinska, K., Czaplinska, M., Swiderski, M., Biesiadka, J. & Legocki, A.B. (2003). *Antonie Van Leeuwenhoek J. Microbiol*. **84**, 115-124.

Weiss, M. S. (2001). *J. Appl. Cryst*. **34**, 130–135.

Table 1 Crystal parameters and data-collection statistics. Values in parentheses correspond to the last resolution shell.

	Native	**Native + SAH**	**Se-Met derivative + SAH**		
			Peak	**Edge**	**Remote**
Space group	$P4_122$ or $P4_322$	$P2_12_12$		$P2_12_12$	
Unit-cell parameters (Å)					
a	48.68	81.01		81.34	
b	48.68	143.30		143.52	
c	141.46	75.85		75.98	
Radiation source	DESY	BESSY		BESSY	
Beamline	EMBL X13	BL 14.2		BL 14.2	
Wavelength (Å)	0.8086	0.9184	0.97968	0.97984	0.95000
Temperature (K)	100	100		100	
No. of molecules in ASU	1	4		4	
Resolution (Å)	2.42 (2.51-2.42)	1.85 (1.92-1.85)		2.00 (2.07-2.00)	
Mosaicity (º)	0.69	0.66		0.45	
Crystal-to-detector distance (mm)	210	190 low res.; 135 high res.		150	
Oscillation (º) /no. of images	0.75/120	1/100 low res.; 0.5/200 high res.	0.75/135	0.75/135	0.75/135
No. of observations	46874 (4020)	366793 (30631)	246696 (17870)	247187 (18400)	250132 (24759)
No. of unique reflections	6954 (638)	72827 (7293)	59766 (4964)	59858 (5111)	59874 (5895)
Completeness (%)	99.6 (97.3)	96.0 (97.8)	98.2 (83.1)	98.5 (85.5)	99.9 (100)
Redundancy	6.7 (6.3)	5.0 (4.2)	4.1 (3.6)	4.1 (3.6)	4.2 (4.2)
R_{merge} [a]	0.087 (0.440)	0.052 (0.347)	0.067 (0.634)	0.054 (0.727)	0.055 (0.423)
$R_{p.i.m.}$ [b]	0.034 (0.176)	0.024 (0.146)	0.037 (0.267)	0.030 (0.306)	0.031 (0.229)
<I/σ(I)>	16.7 (2.1)	24.6 (2.4)	15.7 (2.4)	21.0 (2.3)	21.0 (2.9)

[a] $R_{merge} = \sum_{hkl}\sum_i \left| I_i(hkl) - \langle I(hkl)\rangle \right| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations *i* of reflection *hkl*.

[b] $R_{p.i.m.} = \sum_{hkl}[1/(n_{hkl}-1)]\sum_i | I_i(hkl) - \langle I(hkl)\rangle | / \sum_{hkl}\sum_i I_i(hkl)$ (Weiss, 2001), where n_{hkl} is the number of observations of reflection *hkl*, was calculated in *SCALA* (Evans, 2006) using data processed with *DENZO* (Otwinowski & Minor, 1997).

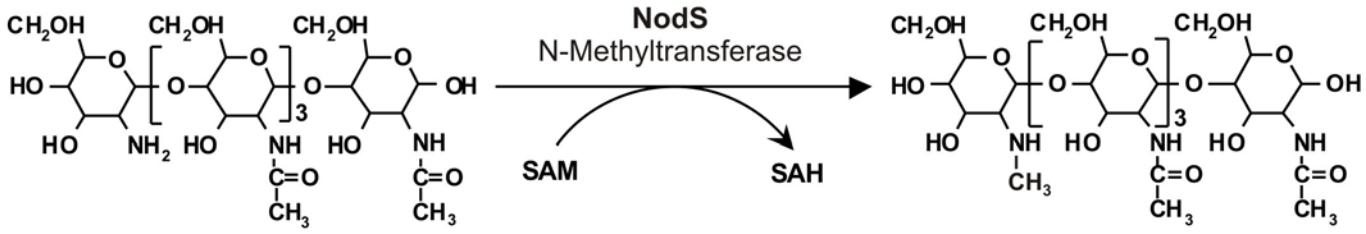

Fig. 1. SAM-dependent methylation step in the biosynthesis of the Nod Factor

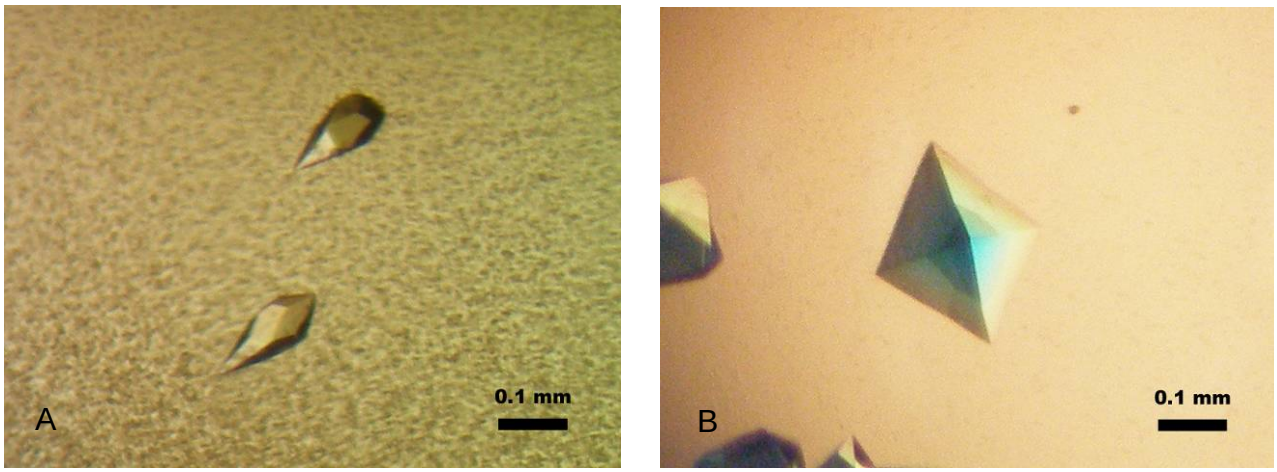

Fig. 2. A: Typical ligand-free NodS crystals (dimensions 0.2 x 0.1 x 0.1 mm). B: A typical single crystal of NodS-SAH complex (dimensions 0.3 x 0.2 x 0.2 mm).

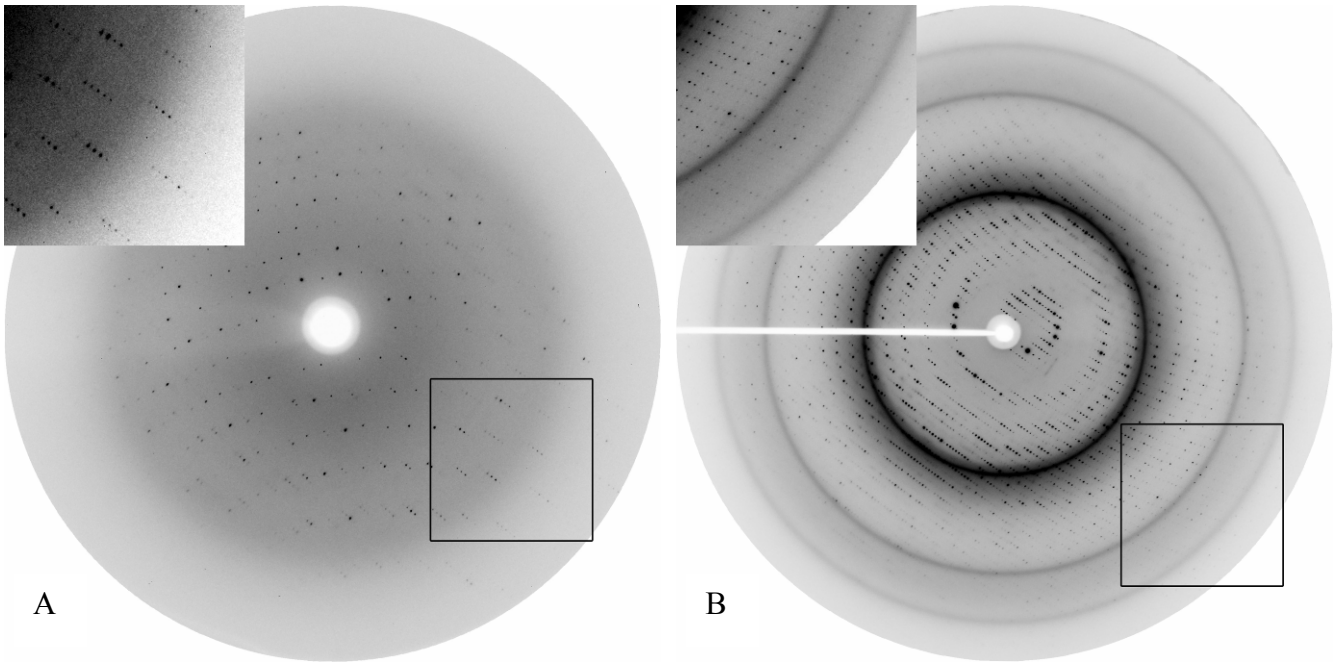

Fig. 3. X-ray diffraction patterns recorded for two different crystals of NodS. A: NodS without ligand (0.75° oscillation), B: NodS in complex with SAH (0.5° oscillation). The edge of the detector (framed, inset) corresponds to a resolution of 2.3 (A) and 1.71 Å (B).