Journal of Structural Biology 164 (2008) 81–87

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Atomic resolution structure of cucurmosin, a novel type 1 ribosome-inactivating protein from the sarcocarp of *Cucurbita moschata*

Xiaomin Hou [a,b], Edward J. Meehan [c], Jieming Xie [d], Mingdong Huang [a,b], Minghuang Chen [a,b,*], Liqing Chen [c,*]

[a] *State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, The Chinese Academy of Sciences, 155 Yang Qiao Xi Lu, Fuzhou 350002, Fujian, China*
[b] *Graduate School of Chinese Academy of Sciences, The Chinese Academy of Sciences, Beijing 10039, China*
[c] *Laboratory for Structural Biology, Department of Chemistry, Graduate Programs of Biotechnology, Chemistry and Materials Science, University of Alabama in Huntsville, Huntsville, AL 35899, USA*
[d] *Fujian Medical University, Fuzhou 350004, China*

ARTICLE INFO

Article history:
Received 27 February 2008
Received in revised form 13 May 2008
Accepted 13 June 2008
Available online 4 July 2008

Keywords:
Ribosome-inactivating protein
Sequence
Structure
Cucurbita moschata
Cytotoxicity
Cancer cells

ABSTRACT

A novel type 1 ribosome-inactivating protein (RIP) designated cucurmosin was isolated from the sarcocarp of *Cucurbita moschata* (pumpkin). Besides rRNA *N*-glycosidase activity, cucurmosin exhibits strong cytotoxicities to three cancer cell lines of both human and murine origins, but low toxicity to normal cells. Plant genomic DNA extracted from the tender leaves was amplified by PCR between primers based on the N-terminal sequence and X-ray sequence of the C-terminal. The complete mature protein sequence was obtained from N-terminal protein sequencing and partial DNA sequencing, confirmed by high resolution crystal structure analysis. The crystal structure of cucurmosin has been determined at 1.04 Å, a resolution that has never been achieved before for any RIP. The structure contains two domains: a large N-terminal domain composed of seven α-helices and eight β-strands, and a smaller C-terminal domain consisting of three α-helices and two β-strands. The high resolution structure established a glycosylation pattern of $GlcNAc_2Man_3Xyl$. Asn225 was identified as a glycosylation site. Residues Tyr70, Tyr109, Glu158 and Arg161 define the active site of cucurmosin as an RNA *N*-glycosidase. The structural basis of cytotoxicity difference between cucurmosin and trichosanthin is discussed.

1. Introduction

The ribosome-inactivating proteins (RIPs) are RNA *N*-glycosidases (Barbieri et al., 1993; Stirpe and Barbieri, 1986) which inactivate ribosomes by site-specifically cleaving the single N–C glycosidic bond between adenine and ribose at A4324 in the 28s rRNA. The depurination of the specific adenine prevents the elongation factor (EF-2) from binding to the 60S subunit, thus RIPs can arrest protein synthesis. Interest in RIPs has arisen from their potential medical and therapeutic applications because several of these proteins have been found to be more toxic to tumor cells than normal cells (Lin et al., 1970). Studies have shown that they can be used as toxic components of immunotoxins targeted at tumor cells (Bolognesi and Polito, 2004; Pastan and Kreitman, 1998), abortifacients (Jin, 1985) and antiviral agents (Au et al., 2000; Wang et al., 2002) in AIDS patients. In addition, it can be used in transgenic reagents in agriculture. Plants transfected with RIP genes exhibit broad-spectrum resistance to viral and fungal infection (Parikh and Tumer, 2004; Wang and Tumer, 2000) in the plant defense system.

Based on the structure of the genes and mature proteins (Van Damme et al., 2001), RIPs can be classified into three types. Type 1 RIPs, such as trichosanthin (Pan et al., 1993), bryodin (Stirpe et al., 1986), α,β-momorcharin (Husain et al., 1994; Yuan et al., 1999) and luffin a, b (Kamenosono et al., 1988), contain a single polypeptide chain and have alkaline isoelectric points. They have potent abilities to inhibit protein synthesis in the cell free system, but are relative non-toxic to the intact cell. Type 2 RIPs, such as ricin (Rutenber et al., 1991) and abrin (Tahirov et al., 1995), consist of two chains, chain A and chain B, linked by disulfide bridges. The A chain possesses the ribosome-inactivating property; the B chain contains a lectin domain which interacts with the cell surface galactosides and facilitates the entry of the A chain into the cytoplasm of the cell. Thus, some, but not all, type 2 RIPs have stronger toxicities than type 1 RIPs because type 1 RIPs can only enter into cells with difficulty, although they are very active towards isolated ribosome. Type 3 RIP is one kind of jasmonate-induced protein, such as JIP60 (Chaudhry et al., 1994) from maize. It consists of an N-terminal domain similar

* Corresponding authors.
E-mail addresses: cmh@fjirsm.ac.cn (M. Chen), chenlq@uah.edu (L. Chen).

1047-8477/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2008.06.011

to other type 1 RIPs and an unrelated C-terminal domain of unknown function.

Cucurmosin possesses an rRNA *N*-glycosidase activity (Chen et al., 2005). Significantly, it can inhibit the proliferation of cancer cells, through the induction of apoptosis (Sun, 2006) and differentiation (Xie et al., 2006) of cancer cells. In this study, we describe the sequencing, cytotoxicity and structural analysis of cucurmosin, a novel type 1 ribosome-inactivating protein, from the sarcocarp of *Cucurbita moschata*.

2. Materials and methods

2.1. Purification, crystallization and data collection

Cucurmosin was isolated and purified to homogeneity from the sarcocarp of *C. moschata* by two-step cation exchange chromatography and crystallized by hanging drop vapor diffusion method as previously described (Chen and Ye, 1999; Chen et al., 2000). The crystals, diffracted to 1.65 Å with in-house rotating anode X-ray source (Shi et al., 2003), diffracted to near atomic resolution of 1.04 Å using synchrotron radiation (APS SER-CAT beamline 22ID) at low temperature (100 K). A 1.04 Å data set was collected and the diffraction data were processed with the program package HKL2000 (Otwinnowski and Minor, 1997). The crystals belong to space group $P2_12_12_1$, with unit-cell parameters $a = 41.76$, $b = 58.78$, $c = 99.52$ Å. Matthews coefficient calculations (Matthews, 1968) show one molecule present in the asymmetric unit and the value of V_m is 1.94 Å/Da, corresponding to a solvent content of 36.7%. The statistics for the data set are summarized in Table 1.

2.2. Structure solution and model refinement

The structure of cucurmosin was solved by the molecular replacement method (AMORE (Navaza, 1994, 2001)) using trichosanthin (Protein Data Bank access code 1TCS) as a search model. Rotational solutions at 8.0–4.0 Å showed one clear peak with the highest correlation coefficient. The translation search was performed with all rotational solutions in the resolution range of 15.0–4.0 Å. The best solution had a correlation coefficient (CC) of 0.568, thrice that of the second (0.175), and an R-factor of 0.438. The refinement was carried out using the program CNS (Brunger et al., 1998) in a resolution range of 50.0–1.04 Å. A total of 5% of the data were randomly selected for R_{free} calculation throughout the whole refinement. A rigid body refinement was carried out first, followed by simulated annealing with energy minimization and restrained individual *B*-factor refinement. The sigma A-weighted $2F_o–F_c$ and $F_o–F_c$ electron density maps were used to guide the model building process. The model was examined and manually rebuilt with the graphic program O (Jones et al., 1991). In the final stage of the refinement, water molecules were added by CNS at locations where electron density was stronger than 3.0σ in sigma A-weighted $F_o–F_c$ maps. The carbohydrates, the polyethylene glycol (PEG) fragments and one phosphate ion were visible at this stage in $2F_o–F_c$ and $F_o–F_c$ electron density maps and were built in. The refinement concluded with the introduction of anisotropic *B*-factors for all non-hydrogen atoms by program Refmac5 in CCP4i (Murshudov et al., 1997, 1999). Data collection and model refinement statistics are listed in Table 1.

Table 1
Data collection and model refinement statistics for cucurmosin

Crystal	Cucurmosin
Space group	$P2_12_12_1$
Cell parameters	$a = 41.76$ Å, $b = 58.78$ Å, $c = 99.52$ Å
Resolutions range (Å)	50–1.04 (1.08–1.04)
Completeness (%)[a]	88.7 (67.8)
Redundancy	7.6 (3.7)
Rmerge (%)[b]	10.5 (35.0)
Unique reflections	106316 (8946)
$I/\sigma(I)$	23.5 (2.0)
R_{work}	0.164 (0.216)
R_{free}	0.172 (0.227)
No. water molecules	258
No. carbohydrates	6
No. of polyethylene glycol fragments	5
No. of phosphate ion	1
R.m.s. deviations from ideal geometry	
Bond lengths (Å)	0.008
Bond angles (°)	1.321

[a] Values in parentheses refer to the highest resolution shell (1.08–1.04) Å.
[b] Rmerge = $\Sigma|I_i - \langle I\rangle|/\Sigma I_i$, where I_i is the intensity of the *i*th observation and $\langle I\rangle$ is the mean intensity of the reflections.

2.3. Cytotoxicity assay

Cucurmosin was tested for its cytotoxicity against human leukemia cells (K562), murine melanoma cells (B16), lung adenocarcinoma cancer cells (A549) and peripheral blood lymphocyte cells using the standard MTT assay, using trichosanthin against human leukemia cells (K562) as positive control. Briefly, 190 μL of cells (1×10^5 cells/mL) were plated in a 96-well plate. Ten microliters of cucurmosin or trichosanthin in various concentrations were added into each well of a 96-well plate. After 72 h of incubation at 37 °C in a humidified 5% CO_2, the cells were added with 20 μL of 5 mg/mL of MTT solution. After 4 h of MTT exposure, the formazan formed was dissolved in 200 μL of dimethy sulfoxide (DMSO). The absorbance was measured at 570 nm.

2.4. DNA sequencing

Total plant DNA was extracted from the tender leaves. Approximately 100 mg of material was frozen and ground to powder in liquid nitrogen. Genomic DNA was further isolated by using the Plant Genomic DNA Extraction Kit (Tiangen, China).

The first 27 residues in N-terminal region were determined by Edman degradation (Chen and Ye, 2003). The X-ray sequence of the C-terminal was originally derived from the high resolution crystal structure. This was necessary to design primers suitable for the PCR experiments.

Based on the N-terminal sequence (NVRFDLS), the oligonucleotide primer (5′-AATGTTAGGTTCGATTTGTC-3′) was used as the upper primer. In this primer, degenerate sites were converted into single nucleotides that were derived from the DNA sequences of homologous protein. From the crystallographic results, X-ray sequence of the C-terminal amino-acids was VVTQNIQ. The oligonucleotide primer (5′-TTGGATGTTYTGGGTTACAAC-3′; were Y = T or C) was used as the lower primer. Weak PCR-product bands with the expected molecular size of about 800 nucleotides were obtained at an annealing temperature of 47 °C. The product was used as template for re-PCR after purification from a gel. Again the product of the re-PCR was purified from a gel and directly used for sequencing. The nucleotide sequence has been submitted to GenBank (EU309692).

3. Results

3.1. Quality of the model

Cucurmosin crystals diffracted to 1.04 Å resolution with synchrotron X-ray radiation. Its structure has been determined and refined at atomic resolution (Table 1). Most residues in the model fit the electron density quite well, except for some residues in loop re-