The three-dimensional structure of nylon hydrolase and its mechanism of nylon-6 hydrolysis

SUPPLEMENTARY DATA ONLINE
Legend for supplementary figures

Fig. S1. Comparison of amino acid sequences of NylC from *Arthrobacter* plasmid pOAD2 (NylC$_{p2}$), *Agromyces* (NylC$_A$) and *Kocuria* (NylC$_K$). Amino acid residues are shown as *one-letter* codes. Fifteen residues differed between NylC$_{p2}$ and NylC$_K$, of which the five residues in *green* boxes are altered in both NylC$_K$ and NylC$_A$. Ten residues in *blue* boxes are unique to NylC$_K$.

Fig. S2. Stereoview of the NylC$_A$ heterodimer shown as a ribbon diagram. The α-subunit and β-subunit in molecule A are colored in *dark green* and *light green*, respectively. The distance (32.6 Å) between Pro260 and Thr267 at C$_\alpha$ is shown.

Fig. S3. Stereoviews of overall structure of NylC$_A$. Superimposition of the NylC$_A$ structure (*green*) with L-aminopeptidase D-Ala-esterase/amidase from *Ochrobactrum anthropi* (DmpA: PDB ID code, 1B65; *magenta*) (**A**), β-peptidyl aminopeptidase from *Sphingosinicella xenopeptidilytica* (BapA: PDB ID code, 3N33; *orange*) (**B**), and ornithine acetyltransferase from *Streptomyces clavuligerus* (PDB ID code, 1VZ6; *yellow*) (**C**). Positions of N-terminus of α-subunit (α-N term), C-terminus of α-subunit (α-C term), N-terminus of β-subunit (β-N term) and C-terminus of β-subunit (β-C term) in NylC$_A$ are shown.

Fig. S4. Multiple three-dimensional alignment of NylC$_A$ and typical N-tn hydrolases. A multiple three-dimensional alignment was performed using secondary structure matching (SSM) of the following proteins (42): DmpA, L-aminopeptidase D-Ala-esterase/amidase (PDB ID code, 1B65); BapA, β-peptidyl aminopeptidase (PDB ID code, 3N33); and OAT, ornithine acetyltransferase (PDB ID code, 1VZ6). The secondary structures are shown as *green* (helix) or *orange* (β-strand) letters. The numbering of DmpA, BapA and OAT is the same as the numbering registered in the protein data bank. Helices and β-strands of NylC$_A$ shown in Fig. 2*A* are illustrated at the top with *green* cylinders (helix) and *orange* arrows (β-strand). The catalytic residue (nucleophile) is in yellow. Amino acid residues affecting the thermostability of NylC$_A$ (Asp36, Ser111, Gly122, Tyr130, Ala137 and Met225) are marked with asterisks. Because the residues of the regions of 261-266 of NylC$_A$ were not able to fit in the electron density map due to poor electron density distribution, they were excluded from the three-dimensional alignment.

Fig. S5. CD spectra at far UV (200-250 nm). *A*. CD measurements were carried out using a J-720WI spectropolarimeter (Jasco, Japan) at 25 °C. ***B*.** A temperature shift from 25°C to 75°C (at 1°C min^{-1}) was performed. CD measurements were conducted at 75°C. A cuvette with a path length of 1 mm was used for all far UV CD measurements. The results are expressed as the mean residue molar ellipticity, [θ], and defined as [θ] = 100 (θ$_{obs}$-θ$_{back}$) l^{-1} c^{-1}, where θ$_{obs}$ is the observed ellipticity in degrees, θ$_{back}$ is the observed ellipticity in degrees in the absence of enzyme (background), c is the molar concentration of the residue, and l is the length of the light path (in centimeters). The protein concentration used in all far UV CD measurements was 0.1 mg ml^{-1}.

Table S1. Sequences of primers used for the construction of NylC mutants.

Mutation	Oligonucleotide	5' Sequence 3'
G111S	FEG111S	GGAGCC**AGC**TACGGGCTCGAGGCGGGC
	REG111S	CCCGTA**GCT**GGCTCCGCCGGCCAGGCA
D122G	FED122G	GGGGTGAGC**GGC**GCGCTCCTGGAACGCCTC
	RED122G	TTCCAGGAGCGC**GCC**GCTCACCCCGGCGCC
H130Y	FEH130Y	CTCGAG**TAT**CGCACCGGCTTCGCCGAGCTC
	REH130Y	GGTGCG**ATA**CTCGAGGCGTTCCAGGAGCGC
L137A	FEL137A	GCCGAG**GCC**CAGCTGGTGTCGTCGGCG
	REL137A	CAGCTG**GGC**CTCGGCGAAGCCGGTGCG
V225M	FV225M	GTGATC**ATG**GACCGCGCGGGCACGGTGGTG
	RV225M	GCGGTC**CAT**GATCACACCGACCGGGTTCGG
D36A	FD36A	GGCAAC**TCC**ACGATCAGCGCGATCGTC
	RD36A	GATCGT**GGA**GTTGCCGGCCTCGGTGACGGG
E263Q	FE263Q	GGCAAC**TGC**ACGATCAGCGCGATCGTC
	RE263Q	GATCGT**GCA**GTTGCCGGCCTCGGTGACGGG

Underlined bases indicate the mutated sites.

Table S2. The root mean square deviation of backbone structure among subunit molecules of NylC$_A$.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
A		0.3	0.29	0.24	0.26	0.22	0.2	0.27	0.23	0.29	0.14	0.24	0.19	0.27	0.15
B	0.3		0.37	.3	0.31	0.28	0.25	0.31	0.34	0.38	0.33	0.29	0.34	0.37	0.32
C	0.29	0.37		0.2	0.19	0.18	0.38	0.22	0.31	0.29	0.26	0.23	0.21	0.23	0.28
D	0.24	0.3	0.2		0.15	0.19	0.33	0.15	0.26	0.27	0.24	0.17	0.19	0.2	0.22
E	0.26	0.31	0.19	0.15		0.18	0.34	0.2	0.27	0.3	0.2	0.21	0.23	0.24	0.25
F	0.22	0.28	0.18	0.19	0.18		0.28	0.23	0.27	0.29	0.24	0.22	0.21	0.24	0.24
G	0.2	0.25	0.38	0.33	0.34	0.28		0.33	0.27	0.37	0.24	0.31	0.3	0.35	0.25
H	0.27	0.31	0.22	0.15	0.2	0.23	0.33		0.27	0.26	0.25	0.17	0.23	0.25	0.24
I	0.23	0.3	0.31	0.26	0.27	0.27	0.27	0.27		0.36	0.24	0.26	0.26	0.3	0.26
J	0.29	0.38	0.29	0.27	0.3	0.29	0.37	0.26	0.36		0.27	0.2	0.26	0.28	0.27
K	0.14	0.33	0.26	0.24	0.26	0.24	0.24	0.25	0.24	0.27		0.23	0.18	0.26	0.17
L	0.24	0.29	0.23	0.17	0.21	0.22	0.31	0.17	0.2	0.2	0.23		0.21	0.22	0.23
M	0.19	0.34	0.21	0.19	0.23	0.21	0.3	0.23	0.26	0.26	0.18	0.21		0.13	0.17
N	0.27	0.37	0.23	0.2	0.24	0.24	0.35	0.25	0.3	0.28	0.26	0.22	0.13		0.24
O	0.15	0.32	0.28	0.22	0.25	0.24	0.25	0.24	0.26	0.27	0.17	0.23	0.17	0.24	

The root mean square deviations (rmsd) (Å) between the fifteen molecules (A - O) at C$_\alpha$ were calculated by secondary structure matching (SSM).

Table S3. Search for proteins with analogous folds using the program DALI

PDB ID	Z	RMSD	LALI	NRES	%ID	PDB description
1B65	26.3	2.8	259	363	20	D-Aminopeptidase
2DRH	25.9	2.7	256	354	21	Hypothetical D-aminopeptidase (361aa long)
3N5I	25.4	3.1	264	371	19	β-Peptidyl aminopeptidase
3N33	25.2	3.6	267	367	18	β-Peptidyl aminopeptidase
3N2W	25.2	3.6	266	367	18	β-Peptidyl aminopeptidase
1VZ8	14.3	3.4	204	376	14	Ornithine acetyl-transferase
1VZ6	14.3	3.4	204	376	14	Ornithine aetyl-transferase
1VZ7	14.1	3.4	204	376	14	Ornithine acetyl-transferase
3IT4	7.3	2.7	136	193	9	Arginine biosynthesis bifunctional protein ArgJ

Z, Z-score (strength of structural similarity in S.D. above expected); RMSD, positional root mean square deviation of superimposed C_α atoms in angstroms; LALI, total number of equivalenced residues; NRES, Number of amino acids in the protein; %ID, percentage of sequence identity.

Table S4. Peptides used for analysis of substrate specificity of NylC$_{p2}$.

Dipeptide	Gly-Gly	Gly-L-Ala	Gly-L-Leu
	Gly-L-Phe	Gly-L-Pro	Gly-L-Val
	L-Ala-L-Ala	D,L-Ala-Gly	D,L-Ala-D,L-Asn
	L-Ala-L-Asp	L-Ala-L-His	D,L-Ala-D,L-Leu
	L-Ala-L-Val	L-Ala-L-Ile	L-Ala-L-Phe
	L-Ala-L-Pro	L-Ala-L-Thr	L-Ala-L-Trp
	L-Ala-L-Tyr	L-Ala-L-Glu	L-Ala-L-Lys
	L-Ala-L-Met	L-Ser-Gly	L-Ser-L-Ala
	L-Val-Gly	L-Val-L-Ala	L-Leu-L-Ala
	L-Leu-Gly	L-Leu-L-Tyr	L-Leu-L-Leu
	L-Pro-Gly	L-Pro-L-Ala	L-Phe-Gly
	L-Phe-L-Ala	L-Trp-L-Ala	L-Trp-Gly
	L-Tyr-Gly	L-Tyr-L-Ala	L-Asp-Gly
	L-Glu-L-Ala	L-Glu-L-Val	L-Met-Gly
	L-Met-L-Ala	L-His-Gly	L-His-L-Ala
	L-His-L-Leu	L-Lys-Gly	
Tripeptide	Gly-Gly-Gly		Gly-Gly-L-Ala
	Gly-Gly-L-His		Gly-Gly-L-Phe
	Gly-L-Pro-L-Ala		Gly-L-Leu-L-Tyr
	Gly-D,L-Leu-D,L-Ala		Gly-L-Ala-L-Ala
	L-Ala-L-Ala-L-Ala		D,L-Ala-Gly-Gly
	D,L-Ala-D,L-Leu-Gly		L-Leu-Gly-Gly
	Gly-L-Phe-L-Ala		L-Glu-L-Val-L-Phe
	L-Ser-L-Gln-Gly		
Tetrapeptide	Gly-Gly-Gly-Gly		
Pentapeptide	Gly-Gly-Gly-Gly-Gly		
	L-Ala-L-Ala-L-Ala-L-Ala-L-Ala		
Hexapeptide	Gly-Gly-Gly-Gly-Gly-Gly		

NylC$_{p2}$ has no detectable activity with the 66 peptides shown in this table (ref. 14, 15).

Positions of amino acid sequence

Enzymes	36	41	50	60	62	111	122	130	137	225	230	231	257	263	354
NyIC$_{p2}$	D	A	M	I	A	G	D	H	L	V	T	V	V	E	G
NyIC$_A$	D	A	M	I	A	S	G	Y	A	M	T	V	V	E	G
NyIC$_K$	A	V	T	V	S	S	G	Y	A	M	G	I	L	Q	A

Fig. S1

Fig. S2

A

α−C term.

α−N term.

β−C term.

β−N term.

B

α−C term.

α−N term.

β−C term.

β−N term.

C

α−C term.

α−N term.

β−C term.

β−N term.

Fig. S3

NylC

6 87

H1 .... H2 * .... S1 .... S2 ...... S3 .... S4

VHALTDIDGGIAVDPAPRLAGPPVFGGPGNDAFDLAPVRSTGREMLRFDFPGVSIGAAHY--EE-----------GPTGATVIHIPAGARTAVDA-----

DmpA

9 79

H1 ...... H2 .... S1 .... S2 .... S3

KPRARDLGLPFTGVTGPYNAITDVDGVGVGFQTII--E-NEPRPGRKRPARSGVTAILPHMQSETPVPVYAGVH

BapA

1 71

H1 ..... H2 .... S1 .... S2 .... S3

GPRARDLGVPFEGTPGALNAITDVAGVEVGHTTVIS--GDGAMVIGKGPYRTGVTIIHPLGKT-SLDGVAAGRA

OAT

8 43

S1 .... S2 ...... S3

TPRGFVVHTAPVGL----------ADDG-RDDFTVLAST------APATVSAV

NylC

88 154

S5 .... S6 .. H3 * .... H4 ... * ... * ... * .. S7

-RGG---AVGLS------GG-Y-----DFNHAICLAG-GASYG--------LEAGAGVSGALLERLE--YRTGFAEAQ--LVSSAVIYDFSARSTA--

DmpA

80 155

H3 .... S4 .... H4 ...... H5 .... H6

RFNG---NGEMT-----GTHWIEDGGYFLGPVVITN-----------THGIGMAHHATVRWMVDRYASTYQTDDF-LWI--MPVVAETYDGALN-DING

BapA

72 145

H3 .... S4 . S5 .... H4 ...... H5 ...... S6

VING---TGEWT-----GMHLVDEVGQFLGPIALTG--------TGNVGLVHQSMMDWSVGKV----PEEALFSRL---LPVVAETLDNRLN-DVFG

OAT

44 115

H1 .... S4 ..... H2 .... H3 S5

F-TRSRFA-GPSVVLCREAV-AD----GQARGVVVLARNANVATGLEGEENAREVREAVARALG--------------LPEGEMLIASTGV-----IG--

---

NylC

155 243

H5 .... S8 S9 .. H6 ... S10 . H7 . S11 ... S12 ...... S13 ...... S14 * . S15

-VYPDKALGRAALEF----AVPGEFPQGRAGAGMSASAGKVDWDRTEITGQGAAFRRLG----DVRILAVVVPNPVG--VIMDRAGTVVRGNYDAQTGVR

DmpA

156 244

H7 .... H8 .... S6 .... S7 ..... H9 ... H10 . H11

FP-VTEADVRKALD----NVASGPVQEGNCGGGTGMITY------GFKGGTGTASRVVEFGGRSFTIGALVQANHGQRDWLTIAGVPVGQHMRDGTPQSQ

BapA

146 233

H6 .... H7 .. S7 .... S8 S9 ...... S10 .... S11 S12 ... S13

HG-LTRDHVFAALDGA---KGGPVAEGNVGGGTGMIAY-----TFKGGIGTSSRVVSAGDTRYTVGVLVQANHGDRNDLRIAGVQIGKEIKGAWPEVN

OAT

116 180

H4 .... H5 .... S6 ..... S7

-RQYPMESIREHLKTLEWPAGEGGFDRAARAIMTTD------------TRPKEVRVSVG----GATLVGIAKGVGMLE--PDMA------------

---

NylC

244 ...... 260 267 330

H8 ..... S16 . H9 ..... S17 .... S18

RHPVFDYQEAFAEQVPPTTISAIVTNV--RMSPVELNQFAKQVHSSM--HRGIQPFHTDMD--GDTLFAVTTDEI---------------DLPTTPGSSRGR

DmpA

245 ...... 250 318

S8 . H12 .... S9 ..... H13

L-----------SIIVVLAT-DLPLMPHQLKRLARRASIGI-GRNGTP-GGNNS-GDIFIAFSTA-NQRPMQHRSAPFLDVEMVN--D-

BapA

234 S14 .... 245 .. 250 S15 . H8 ...... S16

GIVAAGPDAGKP-----SLLIVIAT-DAPLMPHQLERMARRAALGV-GRNGST-AGALS-GEFALAFSTSHVI-PLG--GK-PRLPAIIN-DT------

OAT

181 ...... 250 235

S8 . H6 ..... H7 ...... S9

TLLTFFATDARL-DPAEQDRLFRRVMDRTFNAVSIDT---DTSTSDTAVLFANGLAGE------------V-

---

NylC

331 355

H10

LSV--NATALGAIASEVMWDAVLEAGK

DmpA

319 375

H14 .... H15

----EPLDTVYLAAVDSVEEAVVNAMIAAEDMGGTPFDRLLVQAIDHERLRAVLRQYGRLA

BapA

317 371

H9 .... S17 . S18 . H10

--DSETMNALFRGVVQATEEALVNQLVASETMTGANNAKVYGIPHDQLARIMKARFP

OAT

236 330

H8 .... S10 .. H9 ...... H10 ..... H11 .... S11

-----DAGEFEEALHTAALALVKDIASDGEGAAKLIEVQVTGARDDAQAKRVGKTVVNSPLVKTAVHGCDPNWGRVAMAIGKCSDDTDIDQERVTIRFGE

OAT

331 S12 336 344 H12 ...... S13 ...... S14 .... H13 390

VEVYPPDDALRAAVAEHLRGDEVVIGIDLAIADGAFTVYGCDLTEGYVRLNSE

Fig. S4

A

25°C

B

75°C

Fig. S5