

## CENTRO DE CIÊNCIAS MATEMÁTICAS E DA NATUREZA INSTITUTO DE QUÍMICA

### Processo Seletivo para o Curso de Mestrado do Programa de Pós-Graduação em Bioquímica 2022-1

Artigos e/ou revisões científicas a serem abordados durante a prova de conhecimentos em Bioquímica:

(1) Proteínas: estrutura, função e fracionamento

Na CH, Sathe G, Rosenthal LS, et al. (2020) Development of a novel method for the quantification of tyrosine 39 phosphorylated  $\alpha$ - and  $\beta$ -synuclein in human cerebrospinal fluid. *Clin Proteom.* **17**: 13. doi: <https://doi.org/10.1186/s12014-020-09277-8>

(2) Enzimas: regulação e cinética de Michaelis-Menten

Hu P, Lv J, Fu P, et al. (2013) Enzymatic characterization of an active NDH complex from *Thermosynechococcus elongatus*. *FEBS Letters.* **587**: 2340-2345. doi: <http://dx.doi.org/10.1016/j.febslet.2013.05.040>

(3) Fluxo da informação gênica e técnicas de biologia molecular

Hojny J, Michalkova R, Krkavcova E, et al. (2022) Comprehensive quantitative analysis of alternative splicing variants reveals the HNF1B mRNA splicing pattern in various tumour and non-tumour tissues. *Sci Rep.* **12**: 199. doi: [10.1038/s41598-021-03989-z](https://doi.org/10.1038/s41598-021-03989-z)

(4) Metabolismo: bioenergética, glicólise, ciclo do ácido cítrico e fosforilação oxidativa.

Lima-Silva AE, Fernandes TC, De-Oliveira FR, et al. (2007) Muscle glycogen metabolism during exercise: mechanism of regulation. *Rev Nutr, Campinas.* **20**: 417-429.

RESEARCH

Open Access



# Development of a novel method for the quantification of tyrosine 39 phosphorylated $\alpha$ - and $\beta$ -synuclein in human cerebrospinal fluid

Chan Hyun Na<sup>1,2,3\*</sup>, Gajanan Sathe<sup>3,9</sup>, Liana S. Rosenthal<sup>2</sup>, Abhay R. Moghekar<sup>2</sup>, Valina L. Dawson<sup>1,2,3,5,6</sup>, Ted M. Dawson<sup>1,2,3,6,7\*</sup> and Akhilesh Pandey<sup>3,4,8,10\*</sup>

## Abstract

**Background:** Parkinson's disease (PD) is the second most prevalent neurodegenerative disorder. Biomarkers that can help monitor the progression of PD or response to disease-modifying agents will be invaluable in making appropriate therapeutic decisions. Further, biomarkers that could be used to distinguish PD from other related disorders with PD-like symptoms will be useful for accurate diagnosis and treatment. C-Abl tyrosine kinase is activated in PD resulting in increased phosphorylation of the tyrosine residue at position 39 (Y39) of  $\alpha$ -synuclein ( $\alpha$ -syn) (pY39  $\alpha$ -syn), which contributes to the death of dopaminergic neurons. Because pY39  $\alpha$ -syn may be pathogenic, monitoring pY39  $\alpha$ -syn could allow us to diagnose presymptomatic PD and help monitor disease progression and response to treatment. We sought to investigate if increased phosphorylation of pY39  $\alpha$ -syn can be detected in the cerebrospinal fluid (CSF) of PD patients by targeted mass spectrometry.

**Methods:** Here, we report a two-step enrichment method in which phosphotyrosine peptides were first enriched with an anti-phosphotyrosine antibody followed by a second round of enrichment by titanium dioxide (TiO<sub>2</sub>) beads to detect EGVLPYVGSK sequence derived from tyrosine 39 region of  $\alpha$ - and  $\beta$ -synuclein ( $\alpha\beta$ -syn). Accurate quantification was achieved by adding a synthetic heavy version of pY39  $\alpha\beta$ -syn peptide before enzymatic digestion.

**Results:** Using the developed enrichment methods and optimized parallel reaction monitoring (PRM) assays, we detected pY39  $\alpha\beta$ -syn peptide in human CSF and demonstrated that the ratio of pY39  $\alpha\beta$ -syn to Y39  $\alpha\beta$ -syn was significantly increased in the CSF of patients with PD.

**Conclusions:** We anticipate that this optimized two-step enrichment-based PRM detection method will help monitor c-Abl activation in PD patients and can also be used to quantify other phosphotyrosine peptides of low abundance in biological samples.

**Keywords:** Parkinson's disease,  $\alpha$ -Synuclein,  $\beta$ -Synuclein, Phosphotyrosine, Cerebrospinal fluid, Parallel reaction monitoring

\*Correspondence: chanhyun@jhmi.edu; tdawson@jhmi.edu; pandey.akhilesh@mayo.edu

<sup>1</sup> Neurodegeneration Program, Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

<sup>10</sup> Present Address: Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55902, USA

Full list of author information is available at the end of the article

## Introduction

Although the exact pathogenic mechanism of PD has yet to be established,  $\alpha$ -syn is an important mediator, and moreover, the phosphorylation of  $\alpha$ -syn may contribute to the pathogenesis via increased aggregation and toxicity [1, 2]. Recently, pY39  $\alpha$ -syn has been shown to be



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

closely correlated to disease severity and progression [1, 3, 4]. Thus, pY39  $\alpha$ -syn levels in the brain could potentially serve as a marker for presymptomatic diagnosis, disease progression, and therapeutic response. Since pY39  $\alpha$ -syn is increased in the brain of PD patients, we postulated that this increased phosphorylation could be reflected in the CSF [5].

For targeted detection and quantitation of known proteins or post-translational modifications (PTMs) on those proteins, PRM mass spectrometry (PRM-MS) has been widely used [6, 7]. To quantify pY39  $\alpha$ -syn using PRM-MS, pY39  $\alpha$ -syn peptide, EGVLPYVGSK will be monitored. Although pY39  $\alpha$ -syn peptide is shared between  $\alpha$ -syn and  $\beta$ -syn, we reasoned that the peptide can still serve as a PD biomarker unless the phosphorylation level of  $\beta$ -syn changes to the opposite direction to that of  $\alpha$ -syn in PD. When the abundance of a target protein is too low, the target protein or the derived peptide sometimes can be detected using various enrichment methods such as immunoprecipitation or affinity purification of target proteins or peptides [8–10]. Here, we report a method specifically developed to detect pY39  $\alpha$ -syn peptide in human CSF combining an enrichment method with the PRM-MS approach. Because the abundance of pY39  $\alpha$ -syn in CSF is too low to be detected even with a conventional enrichment method, we developed a two-step enrichment approach using an anti-phosphotyrosine antibody and  $\text{TiO}_2$  beads followed by PRM-MS analysis for detection and quantitation of pY39  $\alpha$ -syn peptide. This method allowed us to establish that the ratio of pY39  $\alpha$ -syn to Y39  $\alpha$ -syn peptides in the CSF can serve as a potential biomarker for the diagnosis and prognosis of PD. Furthermore, this method is applicable to the extremely sensitive detection of other phosphotyrosine peptides as well.

## Methods

### Collection of CSF samples

The CSF specimens were collected from normal pressure hydrocephalus (NPH) patients, PD patients or cognitively normal healthy control individuals evaluated by investigators at the Johns Hopkins Hospital. The CSF samples from NPH patients were used for method optimization. The CSF samples from PD or control individuals were used to compare pY39  $\alpha$ -syn peptide levels between the two groups. The individuals who are cognitively normal or show PD symptoms were diagnosed after extensive clinical and cognitive testing. All PD patients met the UK Brain Bank criteria for PD diagnosis [11]. After the collection of CSF samples by lumbar puncture, the samples were centrifuged for 10 min at  $1500\times g$ , aliquoted, and

stored at  $-80^\circ\text{C}$  within 1 h of acquisition. The demographic and clinical characteristics of the PD patients and control individuals are shown in Additional file 1: Table S1.

### Enrichment of pY39 $\alpha$ -syn peptide only with anti-phosphotyrosine antibody

Approximately  $\sim 5.5$  mg of proteins derived from 9 ml of CSF were lysed in 4 M urea and 50 mM triethylammonium bicarbonate (TEAB) followed by a reduction with 10 mM dithiothreitol for 1 h at room temperature (RT) and alkylation with 30 mM iodoacetamide for 30 min at RT in the dark. The proteins were then digested with an endoproteinase Lys-C (1:100; Wako Chemicals, Richmond, VA) by incubating at RT for 3 h. Subsequently, trypsin digestion was conducted by diluting the urea concentration to 2 M by adding 1 volume of 50 mM TEAB followed by adding sequencing-grade trypsin (1:50; Promega, Madison, WI) and incubating at  $37^\circ\text{C}$  overnight. The peptide samples were desalted with  $\text{C}_{18}$  Sep-Pak (Waters Corporation, Milford, MA) and freeze-dried. The pY39  $\alpha$ -syn endogenous peptide (EGVLPYVGSK) was enriched by performing phosphotyrosine peptide enrichment with PTMScan pY1000 antibody according to the manufacturer's instruction (Cell Signaling Technology, Danvers, MA). Briefly, the  $\sim 2.75$  mg of CSF peptides derived from 9 ml of CSF was reconstituted in 1.4 ml of immunoaffinity purification buffer (IAP, 50 mM MOPS, pH 7.2, 10 mM  $\text{Na}_2\text{HPO}_4$  and 50 mM NaCl). The peptide solution was cleared by centrifugation for 5 min at  $10,000\times g$  at  $4^\circ\text{C}$ , and the supernatant was subject to the phosphotyrosine enrichment. After washing  $40\ \mu\text{l}$  of phosphotyrosine agarose beads three times with PBS, the CSF peptide solution was added to the washed beads followed by incubation at  $4^\circ\text{C}$  for 2 h with rotation. Subsequently, the supernatant was removed, and the beads were washed thrice with 1 ml of IAP buffer and twice with 1 ml of ice-cold water. The bound phosphotyrosine peptides were eluted by adding  $55\ \mu\text{l}$  of 0.15% trifluoroacetic acid (TFA) and incubated at RT for 10 min. After incubation, the tube was centrifuged at  $2000\times g$  for 1 min and the solution was transferred to a new tube. This elution was repeated once again with  $50\ \mu\text{l}$  0.15% TFA. Twenty fmol of synthetic heavy ( $^{13}\text{C}_6$ ,  $^{15}\text{N}_2$ -lysine) pY39  $\alpha$ -syn peptide was added followed by desalting with  $\text{C}_{18}$  StageTip. The eluted peptides were then dried using a SpeedVac followed by reconstitution in  $15\ \mu\text{l}$  of 0.1% formic acid prior to mass spectrometry analysis.

### Enrichment of pY39 $\alpha\beta$ -syn peptide both with anti-phosphotyrosine antibody and TiO<sub>2</sub> beads

For quantification of pY39  $\alpha\beta$ -syn peptides from 1 ml (~0.6 mg of proteins) of CSF samples from PD patients or control samples with both PTMScan pY1000 antibody and TiO<sub>2</sub>, 20 fmol of synthetic heavy pY39  $\alpha\beta$ -syn peptide was added to CSF. CSF proteins were lysed in 4 M urea and 50 mM TEAB followed by a reduction with 10 mM dithiothreitol for 1 h at RT and alkylation with 30 mM iodoacetamide for 30 min at RT in the dark. The proteins were then digested with an endoprotease Lys-C (1:100; Wako Chemicals, Richmond, VA) by incubating at RT for 3 h. Sequentially trypsin digestion was conducted by diluting the urea concentration to 2 M by adding 1 volume of 50 mM TEAB followed by adding sequencing-grade trypsin (1:50; Promega, Madison, WI) and incubating at 37 °C overnight. The peptide samples were desalted with C<sub>18</sub> Sep-Pak (Waters Corporation, Milford, MA) and freeze-dried. The synthetic heavy and endogenous pY39  $\alpha\beta$ -syn peptides were enriched by performing phosphotyrosine peptide enrichment with PTMScan pY1000 antibody according to the manufacturer's instruction with minor modifications (Cell Signaling Technology, Danvers, MA). Briefly, the ~0.3 mg of CSF peptides derived from 1 ml of CSF was reconstituted in 200  $\mu$ l of IAP buffer. The peptide solution was cleared by centrifugation for 5 min at 10,000 $\times$ g at 4 °C and the supernatant was subjected to the phosphotyrosine enrichment. After washing 20  $\mu$ l of phosphotyrosine agarose beads three times with PBS, the CSF peptide solution was added to the washed beads followed by incubation at 4 °C for 2 h with rotation. Subsequently, the supernatant was removed and the beads were washed once with ice-cold water. The bound phosphotyrosine peptides were eluted by adding 55  $\mu$ l of 0.15% TFA and incubated at RT for 10 min. After incubation, the tube was centrifuged at 2000 $\times$ g for 1 min and the solution was transferred to a new tube. This elution was repeated once again with 50  $\mu$ l 0.15% TFA. The eluate was dried using a SpeedVac and the phosphorylated peptides were enriched again using TiO<sub>2</sub> beads as described previously [12]. Briefly, 0.6 mg of TiO<sub>2</sub> beads (Titansphere) resuspended in 40  $\mu$ l of binding buffer (65% acetonitrile (ACN) and 2% TFA) were added to the peptides followed by incubation at RT for 20 min with shaking at 1400 rpm. The peptides were transferred to a C<sub>8</sub> StageTip and centrifuged at 2000 $\times$ g for 2 min. Two hundred  $\mu$ l of the washing buffer (65% ACN and 0.1% TFA) was added and centrifuged at 2000 $\times$ g for 5 min. This washing was repeated once again. The phosphopeptides were eluted by adding 40  $\mu$ l of elution buffer (1% NH<sub>4</sub>OH and 40% ACN) and centrifuging at 200 $\times$ g

for 2 min. The eluted peptides were then dried using a SpeedVac followed by reconstitution in 15  $\mu$ l of 0.1% formic acid prior to mass spectrometry analysis.

### Calculation of the limits of detection and quantification

The limits of detection (LOD) were calculated as:  $LOD = \mu_B + t_{(1-\beta)} (\sigma_B + \sigma_S) / \sqrt{n}$ , where  $\mu_B$  is the estimated mean of blank samples,  $t_{(1-\beta)}$  is 95 percentile of the standard t distribution on f degrees of freedom,  $\sigma_B$  is the standard deviation of the blank samples,  $\sigma_S$  is the standard deviation of the low concentration samples, and n is the number of replicates. The limits of quantification (LOQ) were estimated as 3x LOD [13].

### Detection of Y39 $\alpha\beta$ -syn peptide

To normalize the amount of pY39  $\alpha\beta$ -syn peptide in each sample based on the Y39  $\alpha\beta$ -syn peptide present in each sample, the amount of Y39  $\alpha\beta$ -syn peptide in each sample was also measured. Twenty fmol of heavy (<sup>13</sup>C<sub>6</sub>, <sup>15</sup>N<sub>2</sub>-lysine) Y39  $\alpha\beta$ -syn peptide (EGVLYVGSK\*) for the quantification of the endogenous Y39  $\alpha\beta$ -syn peptide were added to 5  $\mu$ g of CSF peptides followed by desalting with C<sub>18</sub> StageTip and LC-MS/MS analysis.

### LC-MS/MS analysis

The prepared peptides were analyzed on an Orbitrap Fusion Lumos Tribrid Mass Spectrometer coupled to an EASY-nLC 1200 nano-flow liquid chromatography system (Thermo Fisher Scientific). The peptides from each fraction were reconstituted in 15  $\mu$ l of 0.1% formic acid and loaded onto an Acclaim PepMap100 Nano-Trap Column (100  $\mu$ m  $\times$  2 cm, Thermo Fisher Scientific) packed with 5  $\mu$ m C<sub>18</sub> particles at a flow rate of 5  $\mu$ l per min. The flow rate employed was 250 nl/min using a linear gradient of 10% to 35% solvent B (0.1% formic acid in 95% acetonitrile) over 45 minutes on an EASY-Spray column (50 cm  $\times$  75  $\mu$ m ID, Thermo Fisher Scientific) packed 2  $\mu$ m C<sub>18</sub> particles (Thermo Fisher Scientific), which was fitted with an EASY-Spray ion source operated at a voltage of 2.7 kV. Mass spectrometry analysis was completed in a data-dependent manner with a full scan in the mass-to-charge ratio (*m/z*) range of 350 to 1550 followed by targeted MS2. MS1 was measured at a resolution of 120,000 (at *m/z* of 200). MS2 scan was acquired by fragmenting precursor ions using the higher-energy collisional dissociation (HCD) method and detected at a mass resolution of 30,000 (at *m/z* of 200). Automatic gain control was set to 500,000 and 100,000 ions for MS1 and MS2, respectively. The maximum ion injection time for MS1 was set to 100 ms. Maximum ion times for MS2

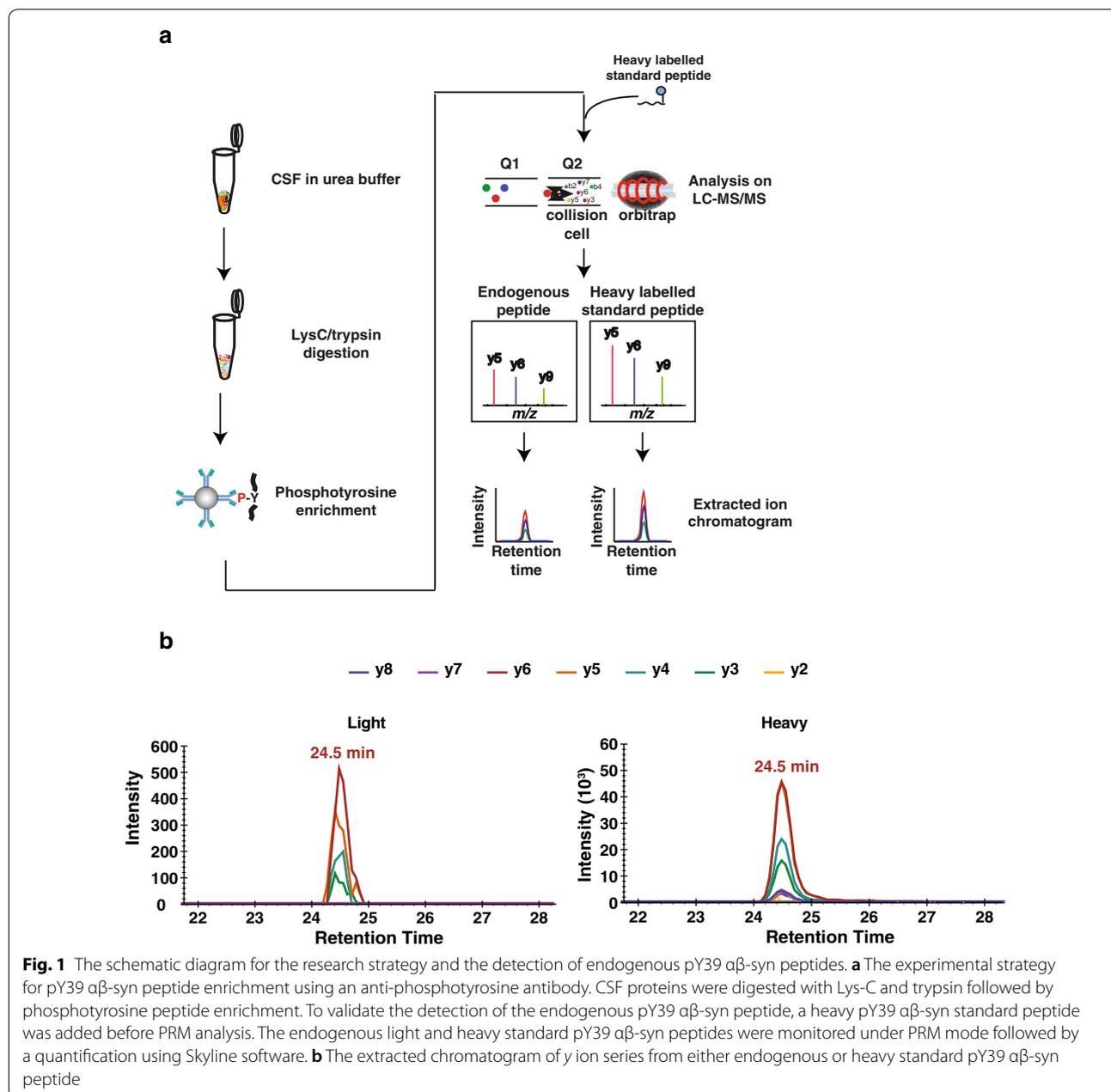
were set to 2500 and 500 ms for pY39 and Y39  $\alpha$ -syn peptides, respectively. HCD normalized collisional energy (NCE) was set to 25, if not specified. The precursor isolation window was set to 1.6  $m/z$ . Internal calibration was carried out using the lock mass option ( $m/z$  445.1200025) from ambient air. For the light and heavy pY39  $\alpha$ -syn peptides,  $m/z$  516.244 and  $m/z$  520.251 were monitored, respectively. For the light and the heavy Y39  $\alpha$ -syn peptides,  $m/z$  476.261 and  $m/z$  480.268 were monitored, respectively.

**Data analysis**

The quantification of relative peptide abundance was performed using Skyline software [14]. The levels of pY39  $\alpha$ -syn were normalized by Y39  $\alpha$ -syn peptide.

**Results**

To detect pY39  $\alpha$ -syn peptide in CSF samples, we initially tried to detect it directly from the digests of CSF samples using PRM but were not able to detect it. Subsequently, we tried an enrichment of  $\alpha$ -syn protein using an anti- $\alpha$ -syn antibody or enrichment of

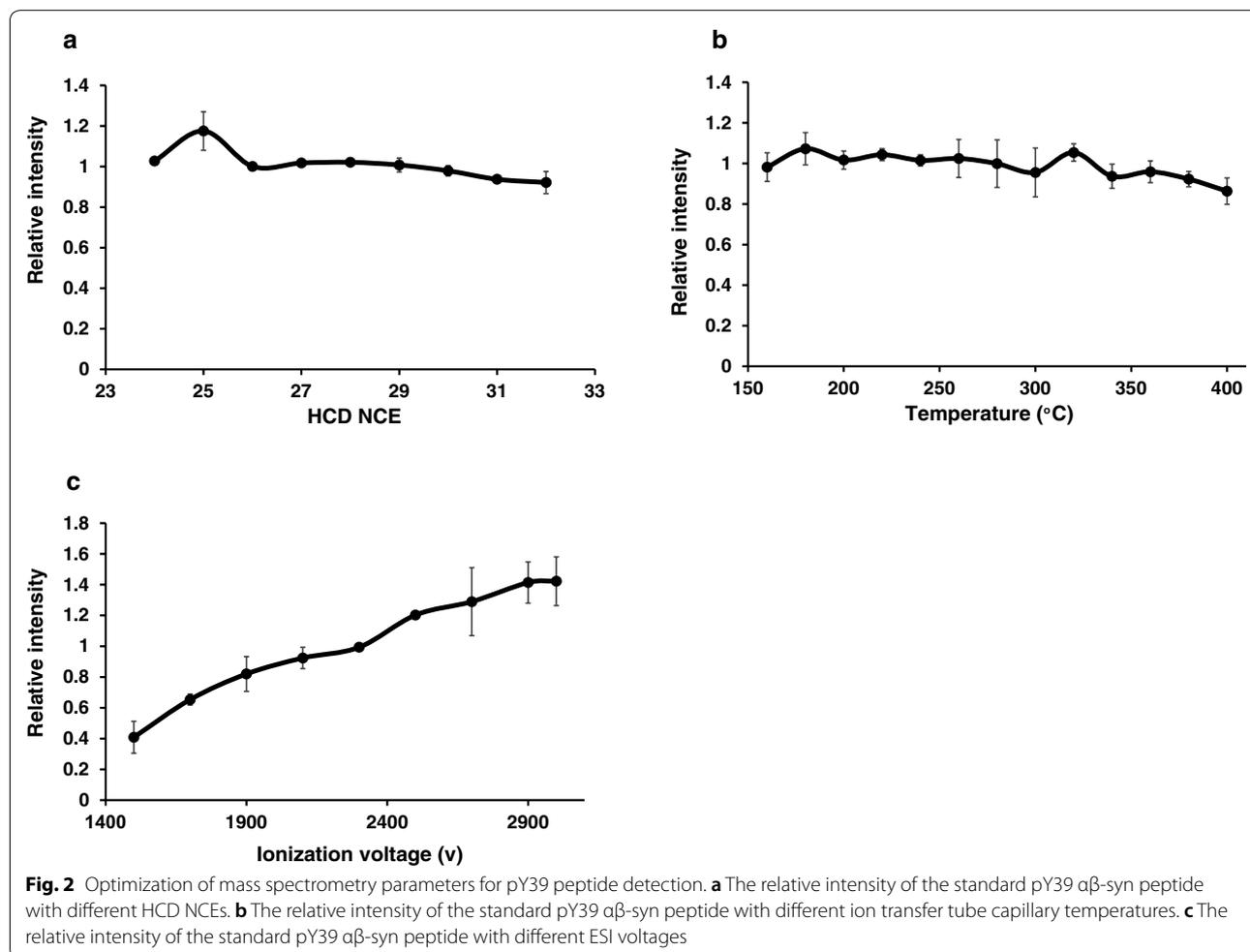


phosphopeptides using  $\text{TiO}_2$  or Immobilized metal affinity chromatography (IMAC). However, none of these enrichment methods were successful in detecting pY39  $\alpha\beta$ -syn peptide. We reasoned that because pY39  $\alpha\beta$ -syn peptide is a tyrosine-phosphorylated peptide, we could further reduce peptide complexity by performing phosphotyrosine peptide enrichment to remove phosphoserine and phosphothreonine peptides, which constitute the majority of phosphopeptides. Using this procedure, we were able to detect the endogenous pY39  $\alpha\beta$ -syn peptide from 9 ml of CSF (Fig. 1a, b). Because 9 ml of CSF is not practical volume in most cases for biomarker detection, we further optimized our methods to reduce the CSF volume required for the analysis.

#### Optimizing the detection of pY39 $\alpha\beta$ -syn peptide

To improve our detection sensitivity, we systematically evaluated the effect of HCD NCE energy, ion transfer capillary temperature and ESI spray voltage on detection.

For the HCD NCE evaluation, the HCD NCE value was increased from 24 to 32. The pY39  $\alpha\beta$ -syn peptide showed the highest intensity at 25 of HCD NCE (Fig. 2a). Next, we evaluated the temperature of the ion transfer capillary increasing it from 160 to 400 °C and pY39  $\alpha\beta$ -syn peptide showed the highest intensity at 180 °C (Fig. 2b). We then evaluated ESI voltage optimization by increasing from 1500 to 3000. pY39  $\alpha\beta$ -syn peptide showed the highest intensity at 3000 V (Fig. 2c). Interestingly, the intensity at 3000 V was >3.5-fold compared to the one observed at 1500 V. Although the HCD NCE and the ion transfer capillary temperature did not change the intensity of pY39  $\alpha\beta$ -syn peptide significantly, the ESI voltage seemed to be critical in increasing the sensitivity. We applied the optimized parameters for the detection of pY39  $\alpha\beta$ -syn peptide except that we opted to use 2700 V instead of 3000 V for preserving column stability. A calibration curve with these optimized parameters permitted us to detect the target peptide at sub-attomole levels. The LOD and LOQ were 0.44 and 1.32 attomoles,



respectively. The coefficient of variations (CV) was calculated as 26.19% at 1 attomole, 4.45% at 10 attomoles, 4.84% at 100 attomoles, 10.03% at 1 femtomole, 4.45% at 10 femtomoles and 8.93% at 100 femtomoles. The average CV of the 6 concentrations was 9.81% (Additional file 2: Figure S1).

#### Development of an enrichment method for pY39 $\alpha\beta$ -syn peptides

To detect the pY39  $\alpha\beta$ -syn peptide from an even smaller volume of CSF, we optimized the enrichment method as well. Even after the enrichment of phosphotyrosine peptides, 80–90% of peptides were still non-phosphorylated ones interfering with the detection of target peptides. Thus, the second step of TiO<sub>2</sub>-based phosphopeptide enrichment should further remove the non-phosphorylated peptides, and thereby, we expected that it would increase the sensitivity of detecting target peptides (Fig. 3a). Thus, we developed a two-step enrichment method by combining the antibody-based phosphotyrosine peptide enrichment method with the TiO<sub>2</sub>-based phosphopeptide enrichment method. To optimize this two-step detection method, we first evaluated the ratio of phosphotyrosine antibody-conjugated agarose beads to the input peptide amount. One femtomole of heavy pY39  $\alpha\beta$ -syn peptide was incubated with various volumes of antibody-conjugated beads ranging from 2.5  $\mu$ l to 80  $\mu$ l. The relative intensity of the heavy pY39  $\alpha\beta$ -syn peptide showed saturation at 20  $\mu$ l of the agarose beads (Fig. 3b). We next optimized the ratio of TiO<sub>2</sub> beads to input peptide amount. For this, 2 femtomoles of heavy pY39  $\alpha\beta$ -syn peptide was incubated with various amounts of TiO<sub>2</sub> beads ranging from 0.1 mg to 3.2 mg. The pY39  $\alpha\beta$ -syn peptide showed the highest intensity at 0.8 mg of TiO<sub>2</sub> beads (Fig. 3c).

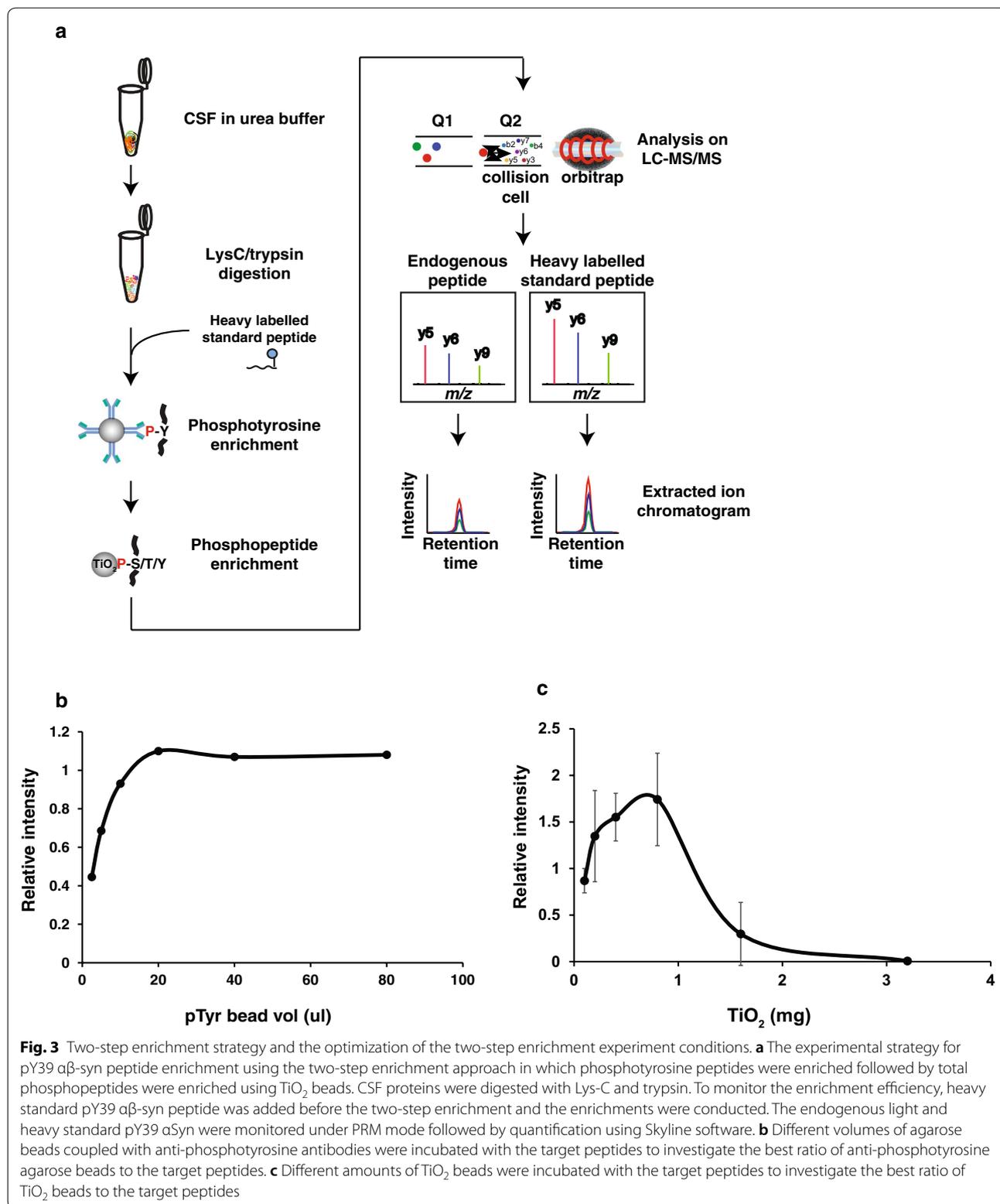
#### Measurement of pY39 $\alpha\beta$ -syn peptide in CSF samples from PD and control individuals

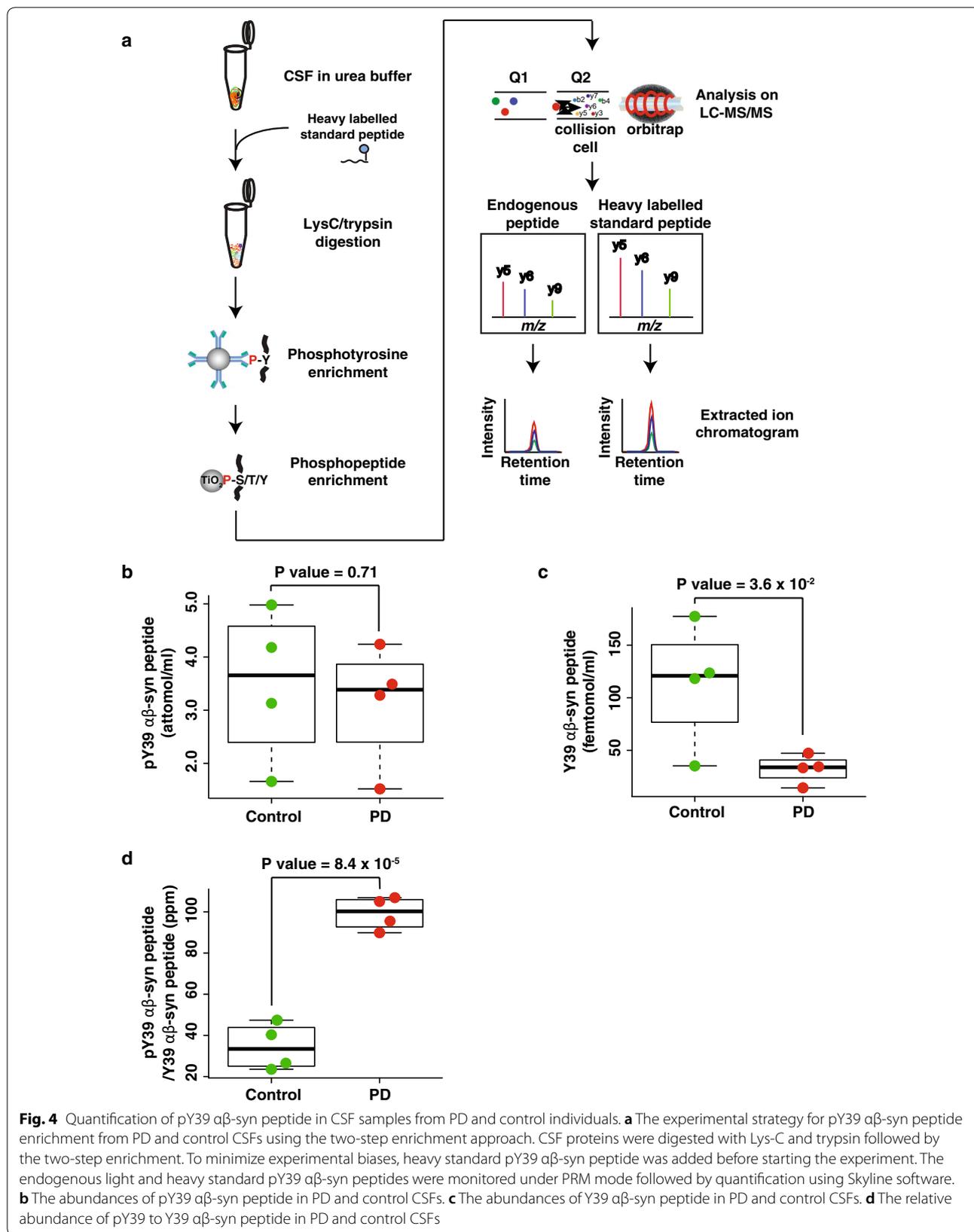
Since the goal of this study was to quantify the endogenous pY39  $\alpha\beta$ -syn peptide in CSF samples from PD patients and control individuals, we applied the optimized mass spectrometry parameters and sample preparation procedure to detect the pY39  $\alpha\beta$ -syn peptides from 4 PD and 4 control individuals as shown in Fig. 4a. To minimize any experimental bias and maximize the accuracy of quantification, synthetic heavy pY39  $\alpha\beta$ -syn peptide was added to all CSF samples at the beginning of the sample preparation step [15]. When pY39  $\alpha\beta$ -syn peptide levels in 1 ml of CSF samples from PD patients were compared to the ones from controls, there was no statistically significant difference (Fig. 4b and Additional

files 3: Table S2, 4). Because it is already known that total  $\alpha$ -syn levels are decreased in PD patients, we postulated that the levels of pY39  $\alpha\beta$ -syn peptide normalized to the ones of Y39  $\alpha\beta$ -syn peptide might help distinguish PD patients from controls. For this, we first measured the Y39  $\alpha\beta$ -syn levels by spiking synthetic heavy Y39  $\alpha\beta$ -syn peptide (Additional file 5: Figure S2). As we expected, Y39  $\alpha\beta$ -syn peptide in CSF from PD patients showed a decreased abundance (Fig. 4c). Most importantly, the ratio of pY39  $\alpha\beta$ -syn to Y39  $\alpha\beta$ -syn peptides in CSF from PD patients showed a noticeable increase (2.5 fold) with statistical significance ( $P$  value =  $8.4 \times 10^{-5}$ ) (Fig. 4d). These results are potentially promising and will have to be validated in a larger cohort.

#### Discussion

In this study, we developed a two-step enrichment method to detect endogenous pY39  $\alpha\beta$ -syn peptide in a sensitive manner from CSF samples using PRM-MS. This approach enabled us to detect the phosphorylated target peptide present at attomole levels per ml of CSF. The enrichment efficiency of phosphotyrosine peptides using the anti-phosphotyrosine antibody-conjugated beads in the first enrichment was usually <15% owing to non-specifically bound peptides even after 5 washes. These non-specifically bound peptides increase the noise in PRM-MS experiments thereby reducing the detection sensitivity. These non-specifically bound peptides can be removed by more extensive washing but it will result in the loss of the target peptide at the same time. For this reason, we chose to remove the non-specifically bound peptides by the second enrichment step using TiO<sub>2</sub> beads instead of washing the beads stringently losing the target peptide. Using this strategy, we were able to improve the detection sensitivity of the target peptide to the level of attomole per ml. While the endogenous pY39  $\alpha\beta$ -syn peptide levels alone did not show a statistically significant difference between PD and control CSFs, the relative abundance of pY39 over Y39  $\alpha\beta$ -syn peptides was strikingly different between the two groups with statistical significance. Our results suggest that the stoichiometry of tyrosine phosphorylation on the residue 39 of  $\alpha\beta$ -syn might be increased in patients with PD, although how much phosphorylation was derived from  $\alpha$ -syn still remains to be elucidated. The development of this method now makes it possible to test the utility of the pY39 over Y39  $\alpha\beta$ -syn peptide ratio as a potential readout for c-Abl activity as well. In addition, this method is broadly applicable to the detection of other phosphotyrosine peptides in biological samples with minor modifications.





## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12014-020-09277-8>.

**Additional file 1: Table S1.** The demographic and clinical characteristics of PD patients and control individuals.

**Additional file 2: Figure S1.** Calibration curve of the heavy synthetic pY39  $\alpha$ -syn peptide in the presence of 10 ng or 2  $\mu$ g of CSF peptides.

**Additional file 3: Table S2.** Quantification results for Y39  $\alpha$ -syn and pY39  $\alpha$ -syn peptides in the CSF of PD patients and control individuals.

**Additional file 4.** Skyline files of the quantification results for pY39  $\alpha$ -syn peptide in the CSF of PD patients and control individuals.

**Additional file 5: Figure S2.** Experimental strategy for the quantification of the Y39  $\alpha$ -syn peptide in CSF.

### Abbreviations

PD: Parkinson's disease;  $\alpha$ -syn:  $\alpha$ -Synuclein;  $\beta$ -syn:  $\beta$ -Synuclein; pY39  $\alpha$ -syn: Phosphorylation of the tyrosine residue at position 39 of  $\alpha$ -syn; pY39  $\alpha$ -syn peptide: EGVLpYVGSK sequence; Y39  $\alpha$ -syn peptide: EGVLYVGSK sequence; TiO<sub>2</sub>: Titanium oxide; PTMs: Post-translational modifications; PRM: Parallel reaction monitoring; NPH: Normal pressure hydrocephalus.

### Acknowledgements

The authors acknowledge the joint participation by the Diana Helis Henry Medical Research Foundation through its direct engagement in the continuous active conduct of medical research in conjunction with The Johns Hopkins Hospital and the Johns Hopkins University School of Medicine and the Foundation's Parkinson's Disease Program H-2013. T.M.D. is the Leonard and Madlyn Abramson Professor in Neurodegenerative Diseases. We also acknowledge the support of the shared instrumentation Grant (S10OD021844).

### Authors' contributions

CHN, TMD, and AP designed research; LSR and ARM collected CSF; CHN and GS performed mass spectrometry analysis. CHN, GS, LSR, ARM, VLD, TMD and AP wrote the manuscript. CHN, TMD and AP supervised research. All authors read and approved the final manuscript.

### Funding

This work was supported by NIH Grants NINDS, P50NS38377 and U01NS082133, U01NS097049 and the Wellcome Trust/DBT India Alliance Margdarshi Fellowship (Grant Number IA/M/15/1/502023 to AP).

### Availability of data and materials

All mass spectrometry data and search results have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD012202 and project name 'Development of a method for the quantification of tyrosine 39 phosphorylated  $\alpha$ -syn in human cerebrospinal fluid' [16]. Reviewers can access the dataset by using 'reviewer43678@ebi.ac.uk' as ID and 'jnapGY1k' as a password.

### Ethics approval and consent to participate

This study was approved by the Johns Hopkins University School of Medicine Institutional Review Board. Informed consent was obtained from all study participants.

### Consent for publication

Not applicable.

### Competing interests

We have no conflict of interest to declare.

### Author details

<sup>1</sup> Neurodegeneration Program, Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>2</sup> Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>3</sup> Diana Helis Henry Medical Research Foundation, New Orleans,

LA 70130, USA. <sup>4</sup> Department of Biological Chemistry, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>5</sup> Department of Physiology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>6</sup> Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>7</sup> Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>8</sup> Manipal Academy of Higher Education (MAHE), Manipal 576104, Karnataka, India. <sup>9</sup> Present Address: Center for Molecular Medicine, National Institute of Mental Health and Neurosciences (NIMHANS), Hosur Road, Bangalore 560 029, India. <sup>10</sup> Present Address: Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55902, USA.

Received: 1 January 2020 Accepted: 18 April 2020

Published online: 04 May 2020

### References

- Brahmachari S, et al. c-Abl and Parkinson's disease: mechanisms and therapeutic potential. *J Parkinsons Dis.* 2017;7(4):589–601.
- Oueslati A. Implication of alpha-synuclein phosphorylation at S129 in synucleinopathies: what have we learned in the last decade? *J Parkinsons Dis.* 2016;6(1):39–51.
- Brahmachari S, et al. Activation of tyrosine kinase c-Abl contributes to alpha-synuclein-induced neurodegeneration. *J Clin Invest.* 2016;126(8):2970–88.
- Burmam BM, et al. Regulation of alpha-synuclein by chaperones in mammalian cells. *Nature.* 2020;577(7788):127–32. <https://doi.org/10.1038/s41586-019-1808-9>.
- Jimenez-Jimenez FJ, et al. Cerebrospinal fluid biochemical studies in patients with Parkinson's disease: toward a potential search for biomarkers for this disease. *Front Cell Neurosci.* 2014;8:369.
- Bourmaud A, Gallien S, Domon B. Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: principle and applications. *Proteomics.* 2016;16(15–16):2146–59.
- Brinkmalm G, et al. A parallel reaction monitoring mass spectrometric method for analysis of potential CSF biomarkers for Alzheimer's disease. *Proteom Clin Appl.* 2018;12:1.
- Preisinger C, et al. Imatinib-dependent tyrosine phosphorylation profiling of Bcr-Abl-positive chronic myeloid leukemia cells. *Leukemia.* 2013;27(3):743–6.
- Ciccimaro E, Hanks SK, Blair IA. Quantification of focal adhesion kinase activation loop phosphorylation as a biomarker of Src activity. *Mol Pharmacol.* 2009;75(3):658–66.
- Razavi M, et al. High precision quantification of human plasma proteins using the automated SISCAPA Immuno-MS workflow. *New Biotechnol.* 2016;33(5):494–502.
- Hughes AJ, et al. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry.* 1992;55(3):181–4.
- Tan H, et al. Refined phosphopeptide enrichment by phosphate additive and the analysis of human brain phosphoproteome. *Proteomics.* 2015;15(2–3):500–7.
- Addona TA, et al. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nat Biotechnol.* 2009;27(7):633–41.
- MacLean B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics.* 2010;26(7):966–8.
- Shuford CM, et al. Peptide production and decay rates affect the quantitative accuracy of protein cleavage isotope dilution mass spectrometry (PC-IDMS). *Mol Cell Proteom.* 2012;11(9):814–23.
- Vizzaino JA, et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 2016;44(D1):D447–56.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



## Enzymatic characterization of an active NDH complex from *Thermosynechococcus elongatus*



Peng Hu<sup>a</sup>, Jing Lv<sup>b</sup>, Pengcheng Fu<sup>b</sup>, Mi Hualing<sup>a,\*</sup>

<sup>a</sup>National Key Laboratory of Plant Molecular Genetics, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, CAS, 300 Fenglin Road, Shanghai 200032, China

<sup>b</sup>Renewable Energy Research Center, China University of Petroleum Beijing, 18 Fuxue Road, Changping, Beijing 102249, China

### ARTICLE INFO

#### Article history:

Received 5 April 2013

Revised 2 May 2013

Accepted 2 May 2013

Available online 27 May 2013

Edited by Richard Cogdell

#### Keywords:

NDH complex

Cyanobacteria

NADPH oxidation

Ferredoxin

Ferredoxin NADP<sup>+</sup> oxidoreductase

Rotenone

### ABSTRACT

**Although type-1 NAD(P)H dehydrogenase (NDH) complex subunit constituents and physiological functions have been reported in plants and cyanobacteria, the biochemical properties of this enzyme are not clear. We used chromatographic isolation to purify and characterize a NADPH-active NDH from the cyanobacterium *Thermosynechococcus elongatus*. Ferredoxin (Fd) and ferredoxin-NADP<sup>+</sup> oxidoreductase (FNR) were co-eluted with NDH, implying the electron donation from NADPH to NDH via the interaction with FNR. We investigated the enzymatic properties of the complex. Furthermore, the activity is competitively inhibited by rotenone, suggesting that it possesses a quinone binding site, similar to mitochondria complex I.**

© 2013 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

### 1. Introduction

Cyanobacterial type-1 NAD(P)H dehydrogenase (NDH) [1] complexes are involved in a variety of bioenergetic reactions, including respiration, cyclic electron transport around photosystem I (PS I) [2] and CO<sub>2</sub> uptake [3]. Of the 14 subunits of complex I in *Escherichia coli*, only 11 *ndh* subunits were identified in cyanobacteria and in chloroplasts [4], three subunits (NuoE, NuoF and NuoG) involved in accepting electrons from NADH in *E. coli* are missing from cyanobacterial and chloroplastic NDH. On other hand, based on proteomics study with several NDH mutants, Battchikova et al. speculate that NdhL-O comprise a domain of unknown function specific for cyanobacteria and chloroplasts, and propose to designate it as the OPS (Oxygenic Photosynthesis-Specific) domain [5]. However, so far, the electron donor and

electron entry for cyanobacterial and chloroplastic NDH are still debatable. Using thylakoid membranes isolated from wild type and NdhB defective mutant M55 of *Synechocystis* PCC 6803, it was demonstrated that NADPH but not NADH donates electrons to the plastoquinone (PQ) pool via NDH [6]. In addition, a reconstitution of NADPH- and ferredoxin (Fd)-dependent cyclic electron flow around PS I has been achieved [6,7]. After then, several NADPH-active NDHs were identified using activity staining for NADPH-NBT oxidoreductase combined with Western blot [8–10]. The activity of a supercomplex of NDH is found to be proportional to cyclic PS I electron flow [11]. To study the biochemical properties of these complex enzymes, active NDH must be purified. However, this presents a considerable technical challenge due to the instability of the complex. Berger et al. [12] first described the isolation of an NDH subcomplex from Triton X-100 treated *Synechocystis* 6803 cell extracts by immunoaffinity chromatography using an NdhK-specific antibody coupled to Protein A Sepharose. The authors monitored the purification with NdhK- and NdhJ-specific antibodies and measurements of NADH-oxidising activity and concluded that the subcomplex was functionally inactive. Later, Matsuo et al. [13] described an active hydrophilic NDH subcomplex of 380 kDa from a CHAPS-treated *Synechocystis* 6803 cell extract. The enzyme appeared to be substrate specific for NADPH but was

**Abbreviations:** NDH, type-1 NAD(P)H dehydrogenase; Fd, ferredoxin; FNR, ferredoxin-NADP<sup>+</sup> oxidoreductase; PBS, phycobilisome protein; HEPES, N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid; DM, n-dodecyl-β-D-maltoside; PS I, photosystem I; DCPIP, 2,6-dichloroindophenol sodium salt dydrate; Q<sub>0</sub>, ubiquinone-0; Q<sub>10</sub>, ubiquinone-10

\* Corresponding author. Fax: +86 21 54924015.

E-mail address: [mihl@sippe.ac.cn](mailto:mihl@sippe.ac.cn) (M. Hualing).

not inhibited by rotenone, a classic inhibitor of Complex I in mitochondria. Deng et al. used conventional chromatographic methods to purify another NADPH-active NDH-1 subcomplex of approximately 230–250 kDa from cells grown under low CO<sub>2</sub> conditions [14]. These authors further showed that a complex containing the hydrophobic subunit NdhA specifically oxidised NADPH [9]. NDH complexes have been isolated from both *Synechocystis* 6803 [15] and *Thermosynechococcus elongatus* [16] using His-tag techniques. A major complex of approximately 460 kDa was isolated from *Synechocystis* 6803, and two complexes of approximately 450 and 490 kDa were isolated in *T. elongatus*, but none of those complexes displayed any NADPH-K<sub>3</sub>Fe(CN)<sub>6</sub> oxidoreductase activity. Nowaczyk et al. [17] isolated NDH-1L complex and identified

two small novel subunits, NdhP and Q from *Thermosynechococcus elongatus*. To date, because all of the isolation approaches described above have been insufficient to obtain either a single purified complex or active NDH, the biochemical properties of NDH remain to be elucidated. Recently, it's reported that knocking out a new identified NdhS in *Arabidopsis thaliana* (also known as CRR31) and in *Synechocystis* 6803 caused inactivation of the activity of NDH pathway [18,19]. According to the structure similarity of NdhS with PsaE, the authors suggested that NdhS might bind with Fd which might be an electron donor for NDH. Here, we report the enzymatic properties of a purified active NDH complex of approximately 170 kDa that contains both hydrophilic and hydrophobic subunits. We found that Fd and Fd-NADP<sup>+</sup> oxidoreductase (FNR) were co-eluted with the active NDH, providing biochemical evidence for interaction of Fd with cyanobacteria NDH and donation of electron from NADPH to NDH via FNR. Furthermore, the activity is competitively inhibited by rotenone.

**Table 1**  
Purification of NDH complex from *T. elongatus*<sup>a</sup>.

Purification step	Total protein (mg)	Total activity <sup>b</sup> (IU)	Specific activity IU (mg protein) <sup>-1</sup>	Purification (fold)	Yield (%)
Cell extracts (DM treated)	1170.00	1815.00	1.552	1	100
Q Sepharose Fast flow	117.47	569.73	4.85	3.13	31.39
Hiload 26/60 Superdex 200	18.21	112.36	6.17	3.98	6.19

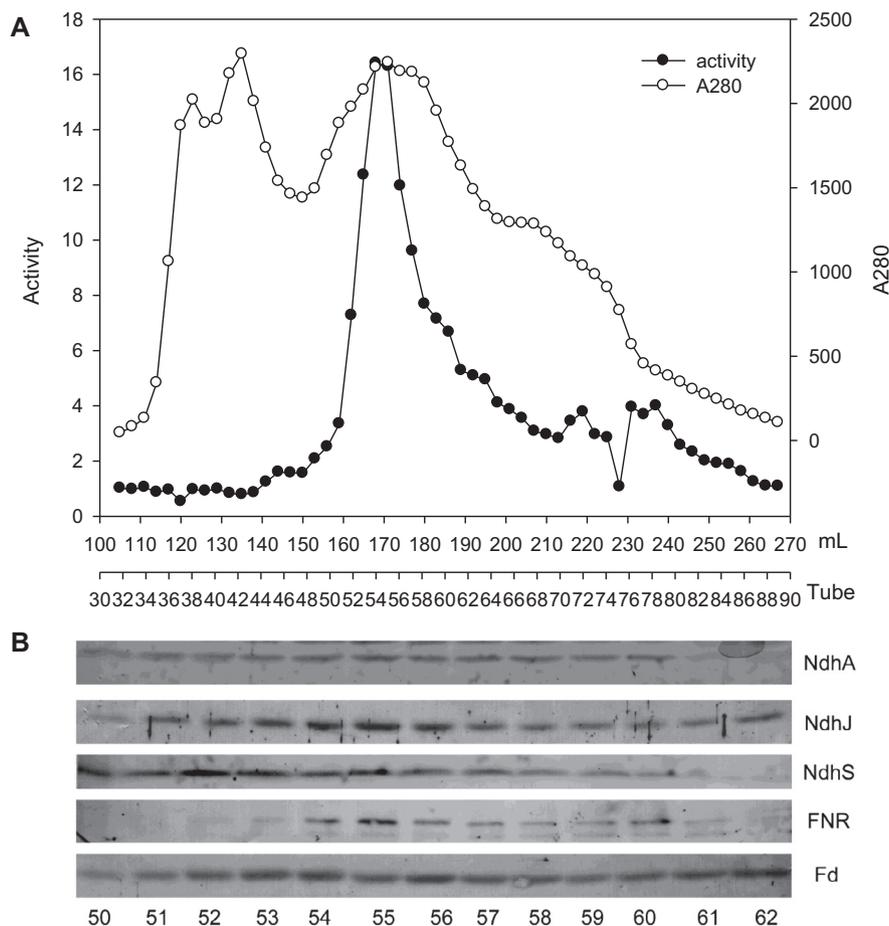
<sup>a</sup> From 10 L of *T. elongatus* cell culture in BG-11 medium.

<sup>b</sup> Activity was determined with 100 μM NADPH, 100 μM K<sub>3</sub>Fe(CN)<sub>6</sub> in 20 mM Tris-Cl, pH 8.0.

## 2. Materials and methods

### 2.1. Organisms and conditions

*Thermosynechococcus elongatus*-BP1 cells were cultured at 50 °C in BG-11 medium [20] buffered with Tris-HCl (5 mM, pH 8.0) bubbled with 2% (v/v) CO<sub>2</sub> in air under continuous illumination by fluorescent lamps (60 μEm<sup>-2</sup> s<sup>-1</sup>).



**Fig. 1.** Size exclusion chromatography. The fractions from the ion exchange step with NADPH oxidation activity were concentrated and applied to Superdex 200 prep grade column. Protein concentration was monitored by absorbance at 280 nm and the activity was detected with a decrease in absorbance at 340 nm using 100 μM NADPH as electron donor and 100 μM K<sub>3</sub>Fe(CN)<sub>6</sub> as an electron acceptor. The same fractions were separated by SDS/PAGE and probed with antibodies specific for NdhA, NdhJ, NdhS, Fd, FNR.

## 2.2. Protein purification

Cells cultured for 4 days ( $A_{730} = 0.6\text{--}0.8$ ) were harvested by centrifugation ( $5000\times g$  for 5 min at  $4\text{ }^{\circ}\text{C}$ ). Cells from 10 l of culture were suspended in 40 ml of medium A (10 mM N-2-hydroxyethyl-piperazine-N'-2-ethanesulfonic acid (HEPES)-NaOH, 5 mM sodium phosphate (pH 7.5), 10 mM  $\text{MgCl}_2$  and 10 mM NaCl) supplemented with 25% glycerol and then stored at  $-80\text{ }^{\circ}\text{C}$ . The thylakoids were isolated as described by Gombos et al. [21], with some modifications. The suspension was mixed with glass beads at a ratio of 1:1 (v/v) and disrupted by 5 pulses of 20 s with a Biospec Bead-beater (Biospec, Japan) followed by a 3 min incubation on ice. The homogenate was centrifuged at  $5000\times g$  for 5 min at  $4\text{ }^{\circ}\text{C}$  to remove unbroken cells and debris, and the supernatant was ultracentrifuged at  $150,000\times g$  for 40 min. The membranes in the pellet were resuspended and solubilised in 1% (w/v) *n*-dodecyl- $\beta$ -D-maltoside (DM) with magnetic stirring on ice for 1 h and then further ultracentrifuged at  $150,000\times g$  for 40 min. The supernatant was immediately subjected to chromatographic separation.

## 2.3. Chromatographic separations

Membrane protein extracts were applied to a Q Sepharose Fast Flow column (Pharmacia, Sweden), equilibrated in buffer A (20 mM Tris-HCl pH 8.0 and 0.05% DM). Elution was performed with buffer B (1 M NaCl, 20 mM Tris-HCl pH 8.0, and 0.05% DM; flow rate 2 ml/min; fraction size 10 ml). Chromatographic fractions with NADPH oxidation activity were concentrated to 2 ml by ultrafiltration (YM 100 membrane; Amicon) at  $4\text{ }^{\circ}\text{C}$ . This extract was loaded onto a HiLoad 26/60 Superdex 200 column (Pharmacia, Sweden) connected to an AKTA system and previously equilibrated in buffer C (150 mM NaCl, 20 mM Tris-HCl pH 8.0, and 0.05% DM). The enzyme was eluted with buffer C (flow rate 1 ml/min; fraction size 1 ml). The fractions that exhibited NADPH oxidation activity were collected for further investigation.

## 2.4. Electrophoresis and protein identification

Native-PAGE was conducted on 7% polyacrylamide gels at  $4\text{ }^{\circ}\text{C}$  and a low constant current of 5 mA as described by Davis [22]. SDS-PAGE was carried out on 15% polyacrylamide gels at room temperature as described by Laemmli [23]. Protein bands were detected by Coomassie Brilliant Blue or  $\text{AgNO}_3$ . Protein bands on SDS gels were transferred to 0.2 mm nitrocellulose membranes. Immunoblot patterns were visualised with the Alkaline Phosphatase (AP) Assay Kit (Bio-Rad).

## 2.5. Kinetics assays

NADPH oxidase activity was measured spectrophotometrically as a decrease in absorbance at 340 nm ( $6.22\text{ mM}^{-1}\text{ cm}^{-1}$ ) at room temperature, as described by Matsuo [13]. The standard assay mixtures containing 1, 10, or 100  $\mu\text{M}$  NADPH and 20 mM Tris-HCl (pH 8.0) were measured with a spectrophotometer (UV-3000, Shimadzu) to determine the effects of various pH values, electron acceptors and inhibitors.

## 2.6. Mass spectrometry

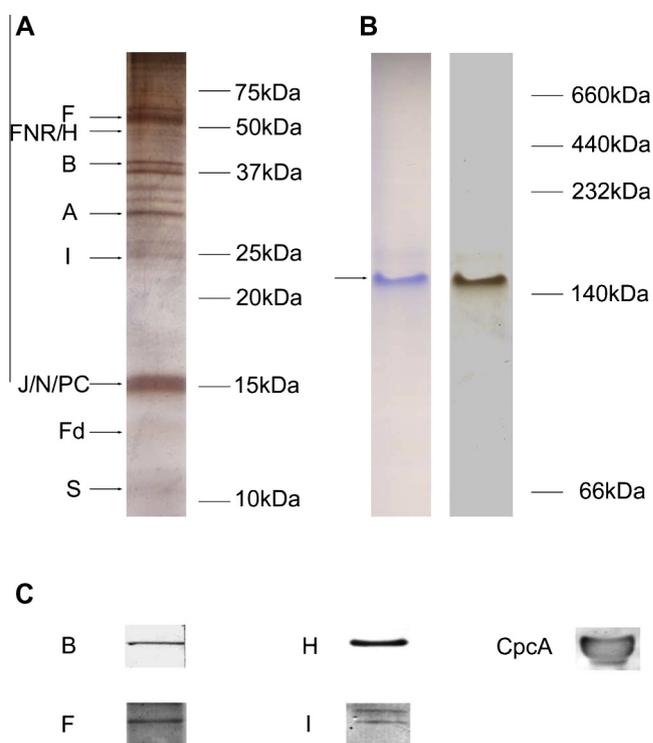
The purified complex was analysed by mass spectroscopy. The samples were prepared according to the manufacturer's instructions. In brief, bands were excised from the native gels, destained, dehydrated, vacuum-dried and incubated overnight with methylated porcine trypsin (Trypsin Gold, Promega). Peptides were analysed with MALDI-TOF using a Voyager-DE-STR mass spectrometer (Thermo Finnigan). Database searches were performed against a

database of *T. elongatus* proteins supplemented with the sequences of common protein contaminants.

## 3. Results

### 3.1. Purification of the NDH complex

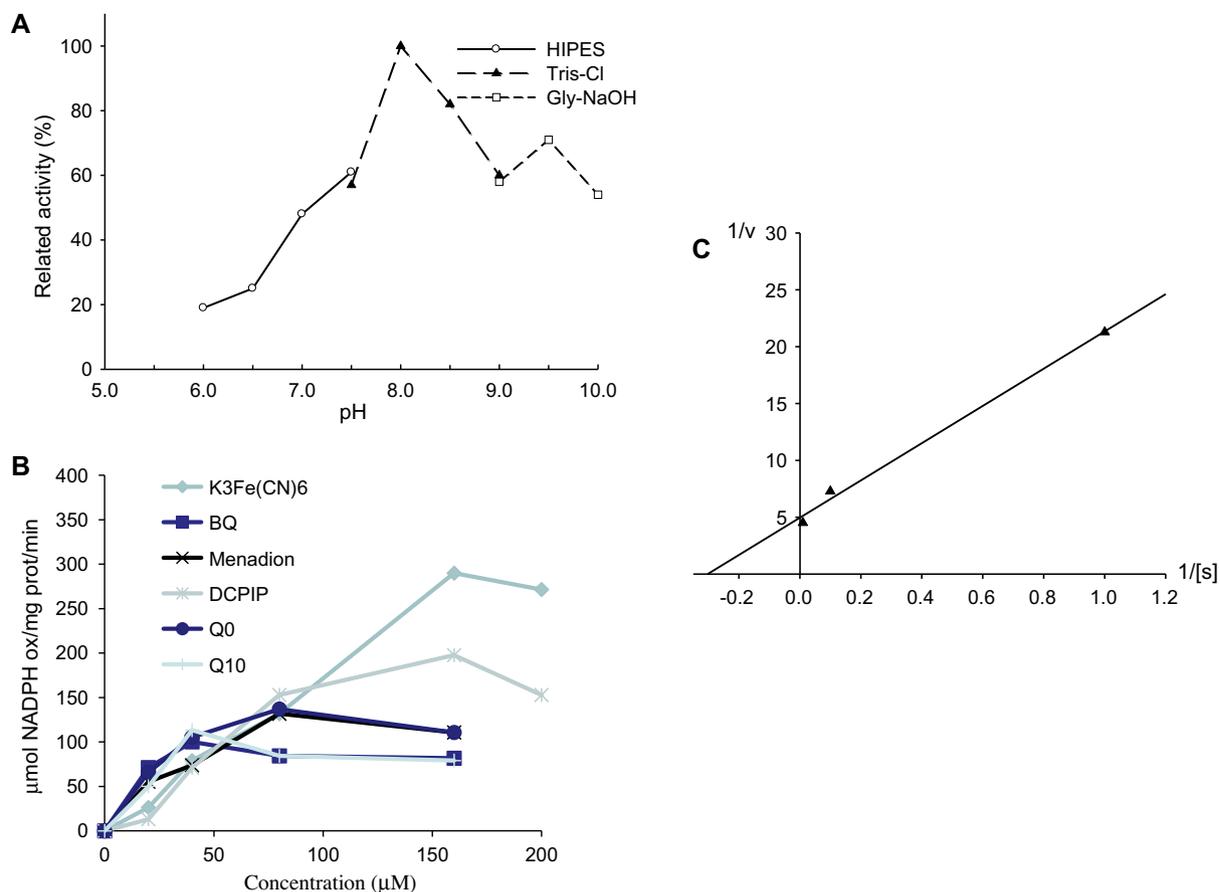
A summary of the purification of the protein complex responsible for the NADPH-oxidation activity in *T. elongatus* is shown in Table 1. Approximately 6 per-cent of the total NADPH-oxidation activity detected in the crude cell extraction was recovered in the fractions eluted from the HiLoad 26/60 Superdex 200 column as a single peak (Fig. 1A). Western blot analysis showed that the content of the NDH subunits NdhA, J, S (Fig. 1B) were proportioned to the NADPH oxidizing activity (Fig. 1), indicating the purified complex is an NADPH-active NDH. In addition, Fd and FNR were co-eluted with the NDH (Fig. 1B), suggesting that the interaction of Fd with the NDH and the electron donation from NADPH might be via FNR. The active protein eluted from the HiLoad 26/60 Superdex 200 column appeared as multiple bands after SDS-PAGE (Fig. 2A) and a single band with an apparent molecular weight about 170 kDa after native PAGE (Fig. 2B). In addition to NdhA, J, S, other NDH subunits NdhB, F, H, and I were identified by Western blot (Fig. 2C), and NdhA, H, and N were identified by mass spectrometric analysis (Table 2) in the complex. Several phycobilisome



**Fig. 2.** Profiling of the NDH complex purified from *T. elongatus*. (A) 15% SDS-PAGE of the NDH complex. (B) Coomassie-stained (left) and  $\text{AgNO}_3$ -stained (right) 7% native PAGE of the NDH complex. (C) Western blot of the NDH complex with antibodies against various NDH subunits.

**Table 2**  
Identification of NDH subunits from purified NDH complex of *T. elongatus* by mass-spectrometry analysis.

Pep count	Identified name
3	tel:tlr1288; NADH dehydrogenase I subunit H
2	tel:tlr1130; NADH dehydrogenase I subunit N
2	tel:tlr0667; NADH dehydrogenase I subunit A



**Fig. 3.** Enzymatic properties of the purified NDH complex. (A) The effect of pH on the NADPH- $K_3Fe(CN)_6$  oxidoreductase activity of NDH complex. The NADH oxidation activity of the NDH complex was measured with  $100 \mu M K_3Fe(CN)_6$  as an electron acceptor at pH values ranging from 6.0 to 10.0 (6.0–7.5 buffered with 20 mM HEPES, 7.5–9.0 buffered with 20 mM Tris-HCl, and 9.0–10.0 buffered with 20 mM Gly-NaOH, respectively). (B) NADPH oxidation activity of the NDH complex with different electron acceptors, buffered with 20 mM Tris-HCl pH 8.0. The following electron acceptors  $K_3Fe(CN)_6$ , DCPIP, menadione (VK<sub>3</sub>), ubiquinone-0 (Q<sub>0</sub>), ubiquinone-10 (Q<sub>10</sub>) were applied. (C) Apparent kinetics of the NADPH- $K_3Fe(CN)_6$  oxidoreductase activity of the NDH complex. NADPH- $K_3Fe(CN)_6$  oxidoreductase activity assays were performed with 0.36 mg/ml purified NDH complex and various NADPH concentrations (1, 10, 100  $\mu M$ ). The  $k_m$  for NADPH and  $V_{max}$  for NADPH- $K_3Fe(CN)_6$  oxidoreductase activity were calculated according to a Lineweaver–Burk plot.

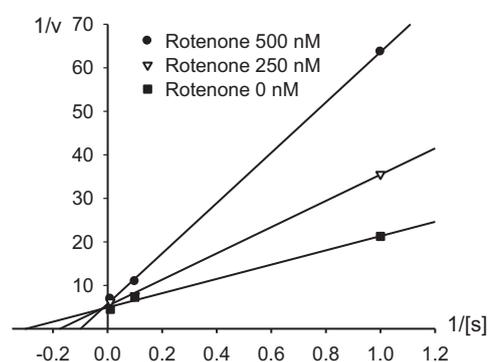
proteins (PBSs) were also detected by Western blot (Fig. 2C), indicating that PBSs may bind to NDH subunits via interaction with FNR to accomplish some biological function.

### 3.2. Optimal pH, electron donors and electron acceptors of the purified active NDH

The dependence of NDH complex activity on pH and electron acceptors was investigated; the results of these investigations are shown in Fig. 3. The optimal pH of the NADPH oxidation of the complex was determined to be 8.0 in Tris-HCl buffer, which is the pH that *T. elongatus* is grown (Fig. 3A). The NDH complex had a higher affinity for PQ analogues than for  $K_3Fe(CN)_6$  and 2,6-dichloroindophenol sodium salt hydrate (DCPIP) (Fig. 3B), implying that PQ is a suitable electron acceptor from NDH. We did not observe any NADH- $K_3Fe(CN)_6$  oxidoreductase activity in the complex with NADH as substrate (Supplementary Fig. 2), indicating that the active NDH specifically oxidises NADPH.

### 3.3. Kinetic properties of the active NDH

Apparent kinetic parameters ( $k_m$  and  $V_{max}$ ) of the NDH complex were calculated using the Lineweaver–Burk plot method. At pH 8.0 and with  $K_3Fe(CN)_6$  as the electron acceptor, experiments with different concentrations of NADPH (1, 10, 100  $\mu M$ ) as substrate showed that the  $k_m$  for NADPH was 3.30  $\mu M$  and the  $V_{max}$  for



**Fig. 4.** The effect of rotenone on NADPH- $K_3Fe(CN)_6$  oxidoreductase activity. NADPH- $K_3Fe(CN)_6$  oxidoreductase activity was assayed as described in Fig. 2C except that 0 nM (control), 250 nM or 500 nM rotenone was added. The Lineweaver–Burk plot demonstrates that rotenone competitively inhibited the activity of the NDH complex.

NADPH- $K_3Fe(CN)_6$  oxidoreductase activity was  $1.01 \mu mol min^{-1} (mg protein)^{-1}$  (Fig. 3C).

### 3.4. The effect of rotenone on NDH activity

Because rotenone, a specific inhibitor of Complex I [24], competitively inhibited the NADPH oxidation activity of a cyanobacterium NDH complex [6], we investigated the effect of rotenone on

the NDH complex purified here. Lineweaver–Burk plot analysis showed that rotenone competitively inhibited the NADPH-oxidation activity of the complex with a  $K_i$  value of 0.26  $\mu\text{M}$  (Fig. 4).

Taken together, the above results strongly suggested that the purified complex responsible for the NADPH-oxidation activity in vitro is identical to the NDH complex in *T. elongatus*.

#### 4. Discussion

There have been reports describing the enzymatic properties of cyanobacterial NDH complexes. Matsuo calculated the  $k_m$  for NADPH of a hydrophilic NDH complex isolated from *Synechocystis* 6803 as 5.1  $\mu\text{M}$  [13], and Deng reported the NADPH  $k_m$  of a hydrophobic NDH complex from *Synechocystis* 6803 as 6.5  $\mu\text{M}$  [14]. Here, we showed that an NDH complex from *T. elongatus* has a  $k_m$  for NADPH of 3.30  $\mu\text{M}$  (Fig. 3C). Although these NDH complexes are different, the values of their  $k_m$  are similar, implying they have common activity domain. We further found that the active NDH had a higher affinity for PQ analogues (Fig. 3B), suggesting that PQ is a suitable electron acceptor for the NDH.

Several distinct NDH subcomplexes, such as NDH-1L, NDH-1MS, NDH-1M, NDH-1S, were discovered by proteomics approaches [25–27]; these complexes are characterised by their distinct NDH subunit compositions. Although the active NDH purified in this report contained NdhA, B, F, H, I, S and N (Figs. 1 and 2), whose were found in NDH-1L, but its molecular weight was lower than that of NDH-1L and it contained FNR and Fd, indicating that it might be another NDH-1 complex. Furthermore, the previously isolated NDH was not inhibited by rotenone (Matsuo et al., 1998); this observation suggests that the previously isolated NDH was a hydrophilic subcomplex that had lost the subunit containing the quinone binding site. In contrast, the NDH purified in this study was sensitive to rotenone (Fig. 4), suggesting that it might contain a quinone binding site similar to that of mitochondria complex I, one of which crystal structure has been resolved very recently [28]. The active NDH also contained NdhS (Fig. 1), the new subunit of NDH identified in *Arabidopsis thaliana* [18] and in *Synechocystis* 6803 [19] responding to the activity of NDH pathway. According to the structure similarity of NdhS with PsaE, the authors suggested that NdhS might bind with Fd which might be an electron donor for NDH. We further detected Fd and FNR co-eluted with the active NDH (Fig. 1), indicating the donation of electron from NADPH to NDH via FNR, as suggested in chloroplast NDH [29]. It has been reported that Fd and FNR form a 1:1 complex, which is stabilized by electrostatic interactions between acidic residues of Fd and basic residues of FNR [30]. Therefore, the low purification fold (Table 1) in this work might be attributable to the dissociation of FNR from the NDH complex during ion exchange isolation and ultrafiltration. However, solid evidence to support the interaction of NDH with Fd and FNR is still lacking in the present study. Such a possibility warrants further investigation. Ndh genes in cyanobacterial genomes show a high homology to those in chloroplast genomes [31]. Therefore, both the NDHs might have similar activity site. It has been indicated that different plastids, derived from proplastids, may have bioenergetic functions and redox chemistry even in the absence of photosynthesis, in which case the NAD(P)H dehydrogenase genes may be retained for the same reason as respiratory complex I in mitochondria of both plant and animal cells [32].

The photosynthetic machinery regulates the distribution of excitation energy between PSI and PSII under fluctuating light conditions. This dynamic and rapid achievement of energy balance is called “state transition” [33,34]. Because PBSs movement is a prerequisite for cyanobacterial state transitions [35–39], “mobile PBSs” are believed to play key roles in allowing state transitions in cyanobacteria. Previous studies have shown that light-induced

state transitions is fully [40] or partially [41] dependent on “mobile PBS” in cyanobacteria. Another study further showed that the movement of PBSs between PSII and PSI affects both cyclic and respiratory electron transport [42]. The inactivation of *ndhB* locked the *Synechocystis* PCC 6803 in state 1 [43]. PBSs such as CpcA were observed in our purified NDH complex through western blot (Fig. 2), probably due to the interaction of the co-eluted FNR (Fig. 1) with PBS [44]. Based on these results, we suggest that PBS and NDH subunits might functionally bind to and be involved in the redistribution of excitation energy in cyanobacteria.

#### Acknowledgements

We thank our colleague Prof. G.Y. Chen and visiting professor T. Ogawa from Japan for fruitful discussion, Dr. W.M. Ma for kindly providing NdhS antibody. This work was supported by funds from the State Key Basic Research and Development Plant 973 [2013CB127005;2011CB200902] and the National Natural Science Foundation of China [31270286].

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.febslet.2013.05.040>.

#### References

- [1] Friedrich, T., Steigmüller, K. and Weiss, H. (1995) The proton-pumping respiratory complex I of bacteria and mitochondria and its homologue in chloroplasts. *FEBS Lett.* 367, 107–111.
- [2] Mi, H.L., Endo, T., Schreiber, U. and Asada, K. (1992) Donation of electrons from cytosolic components to the intersystem chain in the *Cyanobacterium Synechococcus* sp PCC-7002 as determined by the reduction of P700+. *Plant Cell Physiol.* 33, 1099–1105.
- [3] Ogawa, T. (1991) A gene homologous to the subunit-2 gene of NADH dehydrogenase is essential to inorganic carbon transport of *Synechocystis* PCC6803. *Proc. Natl. Acad. Sci. USA* 88, 4275–4279.
- [4] Friedrich, T. and Scheide, D. (2000) The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Lett.* 479, 1–5.
- [5] Battchikova, N., Eisenhut, M. and Aro, E.-M. (2011) Cyanobacterial NDH-1 complexes: novel insights and remaining puzzles. *Biochim. Biophys. Acta Bioenerg.* 1807, 935–944.
- [6] Mi, H.L., Endo, T., Ogawa, T. and Asada, K. (1995) Thylakoid membrane-bound, nadph-specific pyridine-nucleotide dehydrogenase complex mediates cyclic electron-transport in the cyanobacterium *Synechocystis* sp. PCC-6803. *Plant Cell Physiol.* 36, 661–668.
- [7] Mi, H., Endo, T., Schreiber, U., Ogawa, T. and Asada, K. (1994) NAD(P)H dehydrogenase-dependent cyclic electron flow around photosystem-I in the cyanobacterium *Synechocystis* PCC-6803 – a study of dark-starved cells and spheroplasts. *Plant Cell Physiol.* 35, 163–173.
- [8] Mi, H.L., Deng, Y., Tanaka, Y., Hibino, T. and Takabe, T. (2001) Photo-induction of an NADPH dehydrogenase which functions as a mediator of electron transport to the intersystem chain in the cyanobacterium *Synechocystis* PCC6803. *Photosynth. Res.* 70, 167–173.
- [9] Deng, Y., Ye, J.Y. and Mi, H.L. (2003) Effects of low CO<sub>2</sub> on NAD(P)H dehydrogenase, a mediator of cyclic electron transport around photosystem I in the cyanobacterium *Synechocystis* PCC6803. *Plant Cell Physiol.* 44, 534–540.
- [10] Ma, W., Deng, Y., Ogawa, T. and Mi, H. (2006) Active NDH-1 complexes from the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Plant Cell Physiol.* 47, 1432–1436.
- [11] Ma, W., Deng, Y. and Mi, H. (2008) Redox of plastoquinone pool regulates the expression and activity of NADPH dehydrogenase supercomplex in *Synechocystis* sp. strain PCC 6803. *Curr. Microbiol.* 56, 189–193.
- [12] Berger, S., Ellersiek, U., Kinzelt, D. and Steigmüller, K. (1993) Immunopurification of a subcomplex of the nad(p)h-plastoquinone-oxidoreductase from the cyanobacterium *Synechocystis* sp. PCC6803. *FEBS Lett.* 326, 246–250.
- [13] Matsuo, M., Endo, T. and Asada, K. (1998) Properties of the respiratory NAD(P)H dehydrogenase isolated from the cyanobacterium *Synechocystis* PCC6803. *Plant Cell Physiol.* 39, 263–267.
- [14] Deng, Y., Ye, J.Y., Mi, H.L. and Shen, Y.G. (2003) Separation of hydrophobic NAD(P)H dehydrogenase subcomplexes from cyanobacterium *Synechocystis* PCC6803. *Acta Biochim. Biophys. Sin.* 35, 723–727.
- [15] Prommeenate, P., Lennon, A.M., Markert, C., Hippler, M. and Nixon, P.J. (2004) Subunit composition of NDH-1 complexes of *Synechocystis* sp. PCC 6803 –

- identification of two new *ndh* gene products with nuclear-encoded homologues in the chloroplast Ndh complex. *J. Biol. Chem.* 279, 28165–28173.
- [16] Zhang, P.P. et al. (2005) Isolation, subunit composition and interaction of the NDH-1 complexes from *Thermosynechococcus elongatus* BP-1. *Biochem. J.* 390, 513–520.
- [17] Nowaczyk, M.M., Wulffhorst, H., Ryan, C.M., Souda, P., Zhang, H., Cramer, W.A. and Whitelegge, J.P. (2011) NdhP and NdhQ: two novel small subunits of the cyanobacterial NDH-1 complex. *Biochemistry* 50, 1121–1124.
- [18] Yamamoto, H., Peng, L., Fukao, Y. and Shikanai, T. (2011) An Src homology 3 domain-like fold protein forms a ferredoxin binding site for the chloroplast NADH dehydrogenase-like complex in *Arabidopsis*. *Plant Cell* 23, 1480–1493.
- [19] Battchikova, N., Wei, L., Du, L., Bersanini, L., Aro, E.-M. and Ma, W. (2012) Identification of novel Ssl0352 protein (NdhS), essential for efficient operation of cyclic electron transport around photosystem I, in NADPH: plastoquinone oxidoreductase (NDH-1) complexes of *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* 287, 8660.
- [20] Allen, M.M. (1968) Simple conditions for growth of unicellular blue-green algae on plates. *J. Phycol.* 4, 1–4.
- [21] Kombos, Z., Wada, H. and Murata, N. (1994) The recovery of photosynthesis from low-temperature photoinhibition is accelerated by the unsaturation of membrane-lipids – a mechanism of chilling tolerance. *Proc. Natl. Acad. Sci. USA* 91, 8787–8791.
- [22] Davis, B.J. (1964) Disc electrophoresis: 2. Method and application to human serum proteins. *Ann. NY Acad. Sci.* 121, 404–427.
- [23] Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
- [24] Singer, T.P. (1979) Mitochondrial electron-transport inhibitors. *Methods Enzymol.* 55, 454–462.
- [25] Herranen, M., Battchikova, N., Zhang, P.P., Graf, A., Sirpio, S., Paakkarinen, V. and Aro, E.M. (2004) Towards functional proteomics of membrane protein complexes in *Synechocystis* sp. PCC 6803. *Plant Physiol.* 134, 470–481.
- [26] Zhang, P.P., Battchikova, N., Jansen, T., Appel, J., Ogawa, T. and Aro, E.M. (2004) Expression and functional roles of the two distinct NDH-1 complexes and the carbon acquisition complex NdhD3/NdhF3/CupA/Sll1735 in *Synechocystis* sp. PCC 6803. *Plant Cell* 16, 3326–3340.
- [27] Battchikova, N., Zhang, P., Rudd, S., Ogawa, T. and Aro, E.M. (2005) Identification of NdhL and Ssl1690 (NdhO) in NDH-1L and NDH-1M complexes of *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* 280, 2587–2595.
- [28] Baradaran, R., Berrisford, J.M., Minhas, G.S. and Sazanov, L.A. (2013) Crystal structure of the entire respiratory complex I. *Nature* 494, 443–448.
- [29] Guedeney, G., Corneille, S., Cuine, S. and Peltier, G. (1996) Evidence For an association of *ndh B*, *ndh J* gene products and ferredoxin-NADP-reductase as components of a chloroplastic NAD(P)H dehydrogenase complex. *FEBS Lett.* 378, 277–280.
- [30] Depascalis, A.R., Jelesarov, I., Ackermann, F., Koppenol, W.H., Hirasawa, M., Knaff, D.B. and Bosshard, H.R. (1993) Binding of ferredoxin to ferredoxin-NADP+ oxidoreductase – the role of carboxyl groups, electrostatic surface-potential, and molecular dipole-moment. *Protein Sci.* 2, 1126–1135.
- [31] Ellersiek, U. and Steinmuller, K. (1992) Cloning and transcription analysis of the *ndh(A-I-G-E)* gene-cluster and the *ndhD* gene of the cyanobacterium *Synechocystis* sp. PCC6803. *Plant Mol. Biol.* 20, 1097–1110.
- [32] Allen, J.F., de Paula, W.B.M., Puthiyaveetil, S. and Nield, J. (2011) A structural phylogenetic map for chloroplast photosynthesis. *Trends Plant Sci.* 16, 645–655.
- [33] Bonavent, C. and Myers, J. (1969) Fluorescence and oxygen evolution from *Chlorella pyrenoidosa*. *Biochim. Biophys. Acta* 189, 189–366.
- [34] Murata, N. (1969) Control of excitation transfer in photosynthesis 2. Magnesium ion-dependent distribution of excitation energy between 2 pigment systems in spinach chloroplasts. *Biochim. Biophys. Acta* 189, 171–181.
- [35] Allen, J.F., Sanders, C.E. and Holmes, N.G. (1985) Correlation of membrane-protein phosphorylation with excitation-energy distribution in the cyanobacterium *synechococcus*-6301. *FEBS Lett.* 193, 271–275.
- [36] Tsinoremas, N.F., Hubbard, J.A.M., Evans, M.C.W. and Allen, J.F. (1989) P-700 photooxidation in state-1 and state-2 in cyanobacteria upon flash illumination with phycobilin-absorbed and chlorophyll-absorbed light. *FEBS Lett.* 256, 106–110.
- [37] Allen, J.F. (1992) Protein-phosphorylation in regulation of photosynthesis. *Biochim. Biophys. Acta* 1098, 275–335.
- [38] Joshua, S. and Mullineaux, C.W. (2004) Phycobilisome diffusion is required for light-state transitions in cyanobacterial. *Plant Physiol.* 135, 2112–2119.
- [39] Zhang, R., Li, H., Xie, J. and Zhao, J.Q. (2007) Estimation of relative contribution of “mobile phycobilisome” and “energy spillover” in the light-dark induced state transition in *Spirulina platensis*. *Photosynth. Res.* 94, 315–320.
- [40] McConnell, M.D., Koop, R., Vasil'ev, S. and Bruce, D. (2002) Regulation of the distribution of chlorophyll and phycobilin-absorbed excitation energy in cyanobacteria. A structure-based model for the light state transition. *Plant Physiol.* 130, 1201–1212.
- [41] Li, H., Li, D.H., Yang, S.Z., Xie, H. and Zhao, J.Q. (2006) The state transition mechanism - simply depending on light-on and -off in *Spirulina platensis*. *Biochim. Biophys. Acta Bioenerg.* 1757, 1512–1519.
- [42] Ma, W.M., Ogawa, T., Shen, Y.G. and Mi, H.L. (2007) Changes in cyclic and respiratory electron transport by the movement of phycobilisomes in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Biochim. Biophys. Acta Bioenerg.* 1767, 742–749.
- [43] Schreiber, U., Endo, T., Mi, H.L. and Asada, K. (1995) Quenching analysis of chlorophyll fluorescence by the saturation pulse method – particular aspects relating to the study of eukaryotic algae and cyanobacteria. *Plant Cell Physiol.* 36, 873–882.
- [44] van Thor, J.J., Gruters, O.W.M., Matthijs, H.C.P. and Hellingwerf, K.J. (1999) Localization and function of ferredoxin: NADP(+) reductase bound to the phycobilisomes of *Synechocystis*. *EMBO J.* 18, 4128–4136.



OPEN

## Comprehensive quantitative analysis of alternative splicing variants reveals the *HNF1B* mRNA splicing pattern in various tumour and non-tumour tissues

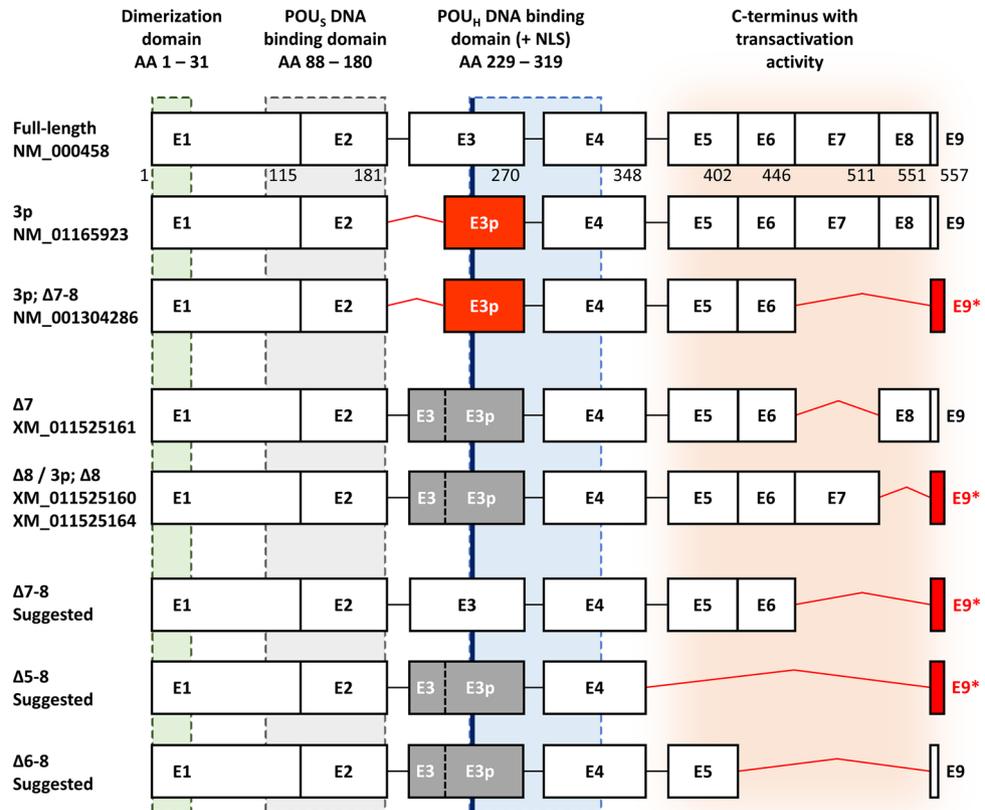
Jan Hojny<sup>1</sup>, Romana Michalkova<sup>1</sup>, Eva Krkavcova<sup>1</sup>, Quang Hiep Bui<sup>1</sup>, Michaela Bartu<sup>1</sup>, Kristyna Nemejcova<sup>1</sup>, Marta Kalousova<sup>2</sup>, Petra Kleiblova<sup>3</sup>, Pavel Dundr<sup>1</sup> & Ivana Struzinska<sup>1</sup>✉

Hepatocyte nuclear factor-1-beta (*HNF1B*) is a transcription factor and putative biomarker of solid tumours. Recently, we have revealed a variety of *HNF1B* mRNA alternative splicing variants (ASVs) with unknown, but potentially regulatory, functions. The aim of our work was to quantify the most common variants and compare their expression in tumour and non-tumour tissues of the large intestine, prostate, and kidney. The *HNF1B* mRNA variants 3p,  $\Delta 7$ ,  $\Delta 7-8$ , and  $\Delta 8$  were expressed across all the analysed tissues in 28.2–33.5%, 1.5–2%, 0.8–1.7%, and 2.3–6.9% of overall *HNF1B* mRNA expression, respectively, and occurred individually or in combination. The quantitative changes of ASVs between tumour and non-tumour tissue were observed for the large intestine (3p,  $\Delta 7-8$ ), prostate (3p), and kidney samples ( $\Delta 7$ ). Decreased expression of the overall *HNF1B* mRNA in the large intestine and prostate cancer samples compared with the corresponding non-tumour samples was observed ( $p = 0.019$  and  $p = 0.047$ , respectively). The decreased mRNA expression correlated with decreased protein expression in large intestine carcinomas ( $p < 0.001$ ). The qualitative and quantitative pattern of the ASVs studied by droplet digital PCR was confirmed by next-generation sequencing, which suggests the significance of the NGS approach for further massive evaluation of the splicing patterns in a variety of genes.

### Abbreviations

CRC	Colorectal carcinoma
ccRCC	Clear cell renal carcinoma
PC	Prostate carcinoma
HGSC	High-grade serous carcinoma
T	Tumour
NT	Non-tumour
p	Deletion/intronization of 5' exon part
$\Delta$	Deletion/intronization of whole exon
ASV	Alternative splicing variant
NGS	Next-generation sequencing
RNA-Seq	RNA next-generation sequencing
ddPCR	Droplet digital PCR
RQN	RNA quality number

<sup>1</sup>Institute of Pathology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Studničkova 2, 12800 Prague 2, Czech Republic. <sup>2</sup>Institute of Medical Biochemistry and Laboratory Diagnostics, First Faculty of Medicine, Charles University and General University Hospital in Prague, Kateřinská 32, 121 08 Prague 2, Czech Republic. <sup>3</sup>Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University and General University Hospital in Prague, Albertov 4, 128 00 Prague 2, Czech Republic. ✉email: ivana.struzinska@vfn.cz



**Figure 1.** Scheme of the proposed *HNF1B* splicing pattern. White boxes represent canonical exons, black lines represent canonical exon-exon junctions. Red boxes represent alternatively spliced exon (E3p) or exon with alternative stop codon (E9\*), red lines represent confirmed alternative exon-exon junctions. Gray boxes represent undetermined variant of exon 3 in suggested transcripts. The lengths of the exons are proportional. Amino acid (AA) numbers are indicated below the exon boxes of full-length transcript. The green, grey, blue, and orange areas illustrate the coding areas for functional domains across the *HNF1B* transcripts. NLS nuclear localization signal (thick blue line). POU<sub>S</sub> POU specific domain. POU<sub>H</sub> POU homeodomain. The scheme of *HNF1B* transcripts is based on previous findings<sup>19</sup> and actual RefSeq database information.

Hepatocyte nuclear factor-1-beta (HNF1B, also known as Transcription Factor-2, TCF2; MIM#189907) is a transcription factor which plays an important role in the regulation of the development of a number of tissues and organs during embryogenesis<sup>1–4</sup>. Apart from its role in differentiation, the HNF1B protein also regulates the expression of multiple genes involved in cell cycle modulation, susceptibility to apoptosis, and response to oxidative stress<sup>5,6</sup>. Its expression in adults was detected in tubule forming epithelial tissues such as kidney and pancreatic exocrine duct tubules, the colon, small intestine, stomach, testes, lungs, liver, prostate etc. (GTEx Portal)<sup>7</sup>. In recent years, a number of studies have suggested that HNF1B may be involved in the tumorigenesis of several types of solid tumours, such as clear cell carcinoma of the ovary, renal cell carcinoma of the kidney, tumours of the gastrointestinal tract, the liver, the pancreas, and the prostate<sup>6,8–14</sup>.

Our previous work showed a loss of the HNF1B protein expression in several solid tumours, and thus suggests that HNF1B may act in a tumour suppressive fashion in colorectal carcinoma (CRC)<sup>15</sup>, clear cell renal carcinoma (ccRCC) and chromophobe renal cell carcinoma<sup>16</sup>, prostate carcinoma (PC)<sup>17</sup>, and high-grade serous carcinoma (HGSC)<sup>18</sup>. The *HNF1B* gene promoter methylation (as a typical factor leading to a loss of protein expression) was observed in 55% PC and 38% HGSC samples. Tumour samples from CRC and ccRCC showed sporadic methylation of the *HNF1B* promoter, as well as the presence of truncating mutations in the *HNF1B* gene.

As a candidate regulatory mechanism of the *HNF1B* gene expression, and due to the incomplete and often contradictory information about *HNF1B* ASVs spectrum, we recently thoroughly investigated the presence and distribution of *HNF1B* mRNA alternative splicing variants (ASVs). The *HNF1B* reference transcript comprises nine coding exons and produces full length 557 amino acid protein (RefSeq NM\_000458). In our previous work, we performed mainly qualitative analyses and described 45 splicing events in a limited number of samples (eight representative samples per analysed tissue pool)<sup>19</sup>. By this approach we have identified predominant and predominant-candidate *HNF1B* ASVs (Fig. 1), which included alternatively spliced exon 3p (exon 3 which lacks 78 bp at 5'; known as variant NM\_01165923.4), Δ7 (deletion/intronization of exon 7; predicted in XM\_011525161.1), Δ7–8 (known in transcript NM\_001304286.2 where was described in combination with 3p variant), Δ8 (predicted in XM\_011525164.1 and XM\_011525160.1), Δ5–8 and Δ6–8 (both not previously described in databases). All the named variants maintain an open reading frame or contain an alternative (not

ASV	Female internal genital tract NT (N = 31) (%)	Kidney NT (N = 31) (%)	Pancreas NT (N = 7) (%)	Prostate NT (N = 35) (%)	Large intestine NT (N = 42) (%)
3p	29.7	28.2	30.2	29.1	33.5
Δ7	1.9	1.7	1.9	1.5	1.7
Δ7–8	1.1	1.0	1.7	0.8	1.2
Δ8	3.2	3.8	6.9	2.3	6.7

**Table 1.** The proportion of alternative splicing variants in selected non-tumour tissues. Each value represents the percentage of the overall *HNF1B* mRNA expression (100%; calculated as the sum of canonical exon 3 and 3p variants). *N* number of samples, *NT* non-tumour tissue.

premature) stop codon. Corresponding protein products were detected and described for 3p and Δ7–8 ASVs (NP\_001159395.1 and NP\_001291215.1, respectively). So far, quantification of ASVs and their co-occurrence in *HNF1B* transcripts have not been fully elucidated. Therefore, *HNF1B* ASVs 3p, Δ7, Δ7–8, Δ5–8, and Δ6–8 ASVs were chosen for subsequent precise quantitative characterization in a wide spectrum of tumour (T) and non-tumour (NT) samples in a number of tissues presented in this work.

The aim of our study was (i) to evaluate the expression of *HNF1B* ASVs and their co-occurrence in the final transcripts in a spectrum of NT samples including the pancreas, large intestine, prostate, kidney and female internal genital tract, and thus to complete the *HNF1B* splicing pattern in non-malignant samples; (ii) to evaluate *HNF1B* overall expression and to quantify the proportion of *HNF1B* ASVs expressed in T sample sets including CRC, ccRCC, PC, and pancreatic carcinoma and their NT counterparts; (iii) to show the usability of the capture RNA next-generation sequencing (RNA-Seq) approach for the detection and quantification of ASVs. The precise characterization of the *HNF1B* splicing pattern in T and NT tissues may contribute to a better general understanding of the *HNF1B* gene expression and the role of *HNF1B* ASVs in tumorigenesis.

## Results

**Expression of ASVs 3p, Δ7, Δ7–8, and Δ8 was detected in all analysed NT tissue samples.** The droplet digital PCR (ddPCR) analysis of NT samples revealed the expression of the 3p, Δ7, Δ7–8, and Δ8 ASVs in all 146 NT samples in the subset of samples with sufficient *HNF1B* mRNA expression (see methods). The proportion of *HNF1B* ASVs is similar among analysed tissue sets (Table 1), although the overall *HNF1B* expression level differs in different tissues<sup>7</sup>.

Variants Δ5–8 and Δ6–8 were detected during the initial analysis steps only in a portion of NT kidney samples (15/48 T and 31/48 NT). The measured expression levels were very low and close to the detection limit of the method used (< 0.5% of overall *HNF1B* expression). Considering the highest overall *HNF1B* mRNA expression in the NT kidney tissue compared to other tissues<sup>7</sup>, these variants were not analysed in other sample sets, and we considered them as minor *HNF1B* ASVs.

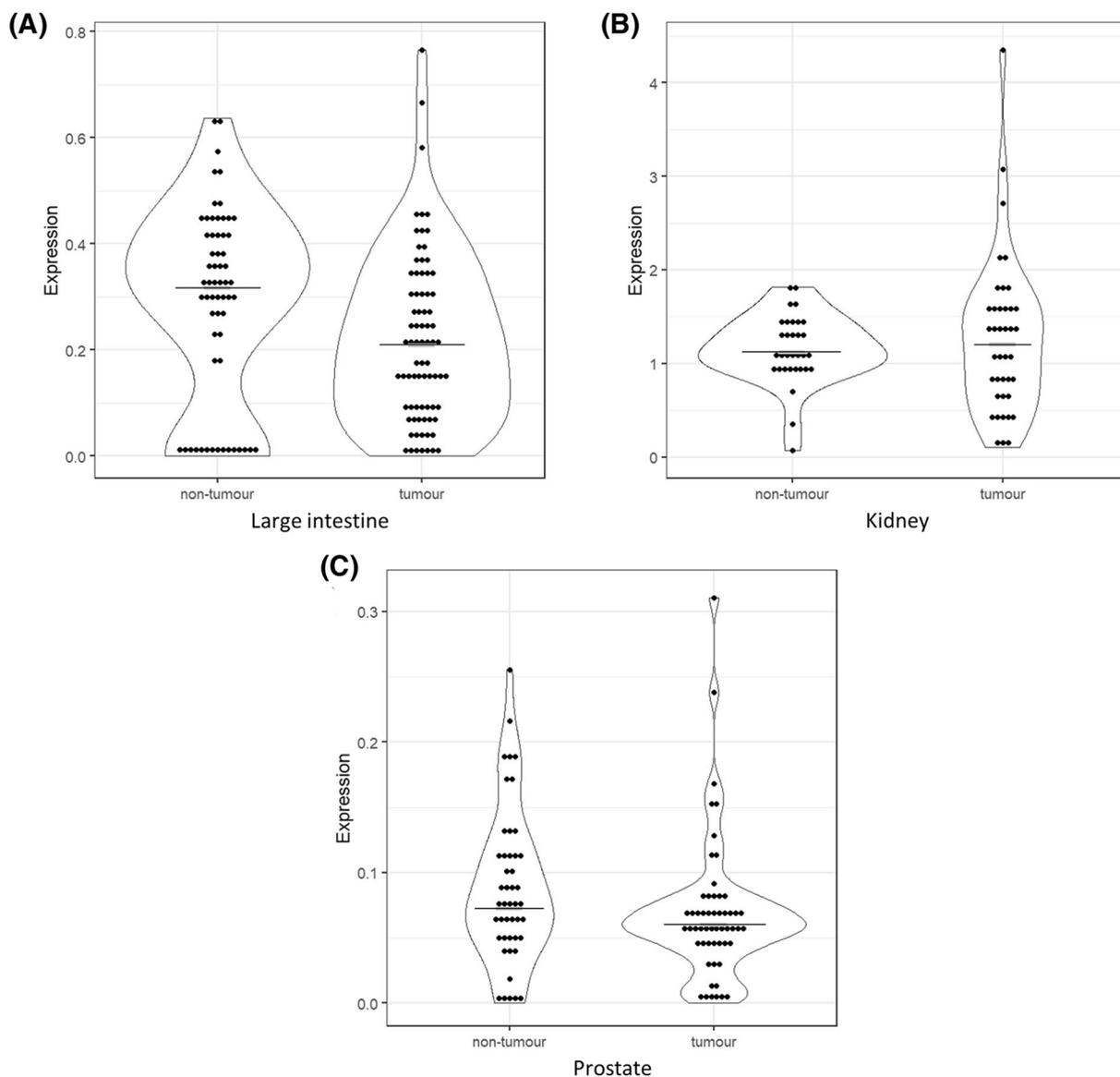
**ASVs Δ7, Δ7–8 or Δ8 exist also in combination with the ASV 3p in *HNF1B* transcripts.** In a representative pool of NT kidney samples (see methods) PCR analysis of the *HNF1B* transcripts containing Δ7, Δ7–8 or Δ8 ASV was performed. In the PCR reactions unique primers for each ASV were used together with forward primer in exon 2, which allowed us to determine the variant of exon 3. The analysis confirmed the existence of transcripts containing each of these ASVs separately (with canonical exon 3), as well as in combination with the 3p ASV (Fig. S1).

**Reduced overall *HNF1B* mRNA expression was detected in colorectal and prostate cancer and correlates with low protein expression in colorectal cancer.** A comparison of the overall *HNF1B* expression between T and NT tissue samples was successfully performed in 78 T and 61 NT large intestine tissue samples, 40 T and 32 NT kidney samples, and 55 T and 46 NT prostate samples. A decreased overall expression of *HNF1B* was observed in the large intestine T samples in comparison to the NT counterparts ( $p = 0.019$ ; Fig. 2A), and in prostate T samples in comparison to the NT samples ( $p = 0.047$ , Fig. 2C). There were no statistically significant differences in the expression observed in the T and NT kidney samples (Fig. 2B).

Paired comparisons of the overall *HNF1B* expression within the subsets of matched T and NT samples were performed in 61 large intestine tissue samples, 31 kidney tissue samples, and 46 prostate tissue samples. This statistical approach showed the same results compared to non-paired testing. The complete results are listed in Table 2A.

We further correlated the mRNA expression with protein expression in 78 CRC. The immunohistochemical protein expression (H-score) of the same CRC sample set was assessed previously by our group<sup>15</sup>. The comparison revealed a weak positive correlation ( $R = 0.39$ ,  $F = 13.8$ ,  $df = 1.76$ ,  $p < 0.001$ ; Fig. 3). A similar analysis was published previously for the same sample set of prostate carcinomas, where no correlation between mRNA and protein expression was observed<sup>17</sup>.

**Differences between the relative expression of the *HNF1B* ASVs in tumour and non-tumour samples.** The analysis of *HNF1B* ASVs expression was successfully performed in the subset of samples with sufficient *HNF1B* expression (see methods). The statistical evaluation was performed in 57 T and 42 NT tissue samples of large intestine, 37 T and 31 NT kidney samples, and 42 T and 35 NT prostate samples. The rela-



**Figure 2.** *HNF1B* mRNA overall expression levels in the tumour and non-tumour samples. (A) large intestine (NT = 61 samples; T = 78 samples;  $p = 0.019$ ); (B) kidney (NT = 32; T = 40;  $p = 0.968$ ); (C) prostate tissue (NT = 46; T = 55;  $p = 0.047$ ). Data is visualized as violin plots. Each dot represents one sample. Expression is relative to *POLR2A* (*POLR2A* expression = 1). Black line represents median.

tive expression of the 3p variant was significantly decreased in T compared with NT tissues of the large intestine samples (median T = 31.6%; NT = 33.5%;  $p = 0.018$ ; Fig. 4A) and the prostate samples (median T = 26.5%; NT = 29.1%;  $p < 0.001$ ; Fig. 4C). On the contrary, an increased expression of the  $\Delta 7$  variant was detected in the kidney tissue T samples when compared with the kidney NT samples (median T = 2.2%; NT = 1.7%;  $p = 0.037$ , Fig. 4B). An increased expression of the  $\Delta 7-8$  variant was detected in T samples in comparison to NT samples in the prostate tissue (median T = 1.0%; NT = 0.8%;  $p = 0.028$ , Fig. 4C), and the large intestine samples (median T = 1.9%; NT = 1.2%;  $p = 0.001$ , Fig. 4A). The expression of the alternative variant  $\Delta 8$  did not show any statistically significant differences in the tissues analysed. The complete results are provided in Table 2B.

Paired comparisons of ASVs within the matched T and NT samples were performed in 33 large intestine tissue samples, 28 prostate tissue samples, and 28 kidney tissue samples. In accordance with the non-paired testing, significant differences were observed for the 3p ASV in the large intestine and prostate paired samples, the  $\Delta 7$  ASV in the kidney paired samples, and the  $\Delta 7-8$  ASV in the large intestine. Although the comparison between statistical approaches of individual ASVs showed the same trend, no statistically significant differences were observed in the case of  $\Delta 7-8$  ASV in the prostate and kidney paired samples. The complete results are provided in Table 2B and Fig. S2.

The *HNF1B* DNA mutation analysis of this sample set was done and published previously<sup>15-17</sup> and did not reveal any variant located in canonical or potential cryptic splice site of T or NT samples in the coding *HNF1B* regions with flanking intronic sequences (+ – 20 bp). Overall, few different variants were detected in our dataset. All these variants were tested for potential splice effect by in-silico tools with any potential splice predictions.

<b>A)</b>							
overall	Large intestine	N	Kidney	N	Prostate	N	
<i>HNF1B</i>							
	<b>1</b>	<b>Z = -2.35, p = 0.019</b>	78 (T) 61 (NT)	Z = 0.04, p = 0.968	40 (T) 32 (NT)	<b>Z = -1.99, p = 0.047</b>	55 (T) 46 (NT)
	<b>2</b>	<b>Z = -2.39, p = 0.017</b>	61	Z = 0.37, p = 0.709	31	<b>Z = -2.10, p = 0.035</b>	46
<b>B)</b>							
<i>HNF1B</i>	Large intestine	N	Kidney	N	Prostate	N	
ASVs							
3p	<b>1</b>	<b>Z = -2.35, p = 0.018</b>	57 (T) 42 (NT)	Z = 0.80, p = 0.423	37 (T) 31 (NT)	<b>Z = -3.34, p &lt; 0.001</b>	42 (T) 35 (NT)
	<b>2</b>	<b>Z = -3.24, p = 0.001</b>	33	Z = 1.16, p = 0.245	28	<b>Z = -2.21, p = 0.027</b>	28
$\Delta 7$	<b>1</b>	Z = 1.56, p = 0.118	57 (T) 42 (NT)	<b>Z = 2.08, p = 0.037</b>	37 (T) 31 (NT)	Z = -0.76, p = 0.446	42 (T) 35 (NT)
	<b>2</b>	Z = 1.74, p = 0.081	33	<b>Z = 2.23, p = 0.026</b>	28	Z = -1.23, p = 0.218	28
$\Delta 7-8$	<b>1</b>	<b>Z = 3.28, p = 0.001</b>	57 (T) 42 (NT)	Z = 1.69, p = 0.089	37 (T) 31 (NT)	<b>Z = 2.19, p = 0.028</b>	42 (T) 35 (NT)
	<b>2</b>	<b>Z = 2.15, p = 0.031</b>	33	<b>Z = 2.28, p = 0.023</b>	28	Z = 1.29, p = 0.194	28
$\Delta 8$	<b>1</b>	Z = -0.87, p = 0.386	57 (T) 42 (NT)	Z = 0.72, p = 0.467	37 (T) 31 (NT)	Z = 1.37, p = 0.169	42 (T) 35 (NT)
	<b>2</b>	Z = -1.17, p = 0.242	33	Z = 0.73, p = 0.466	28	Z = 0.11, p = 0.909	28

**A)** overall *HNF1B* expression; **B)** expression of ASVs; 1 - Mann-Whitney U test was applied for independent samples; 2 - Wilcoxon matched-pairs signed-ranks test was used for matched T and NT samples. Statistically significant results are indicated in bold. N - number of samples; T - tumour tissue; NT - non-tumour tissue; Z - standardized score, critical value for a 95% confidence interval; p - level of marginal significance within statistical test.

**Table 2.** Comparison of overall expression or individual ASVs' expression between tumour and non-tumour samples in different tissues. (A) overall *HNF1B* expression; (B) expression of ASVs; 1—Mann–Whitney U test was applied for independent samples; 2—Wilcoxon matched-pairs signed-ranks test was used for matched T and NT samples. Statistically significant results are indicated in bold. N number of samples; T tumour tissue, NT non-tumour tissue, Z standardized score, critical value for a 95% confidence interval, p level of marginal significance within statistical test.

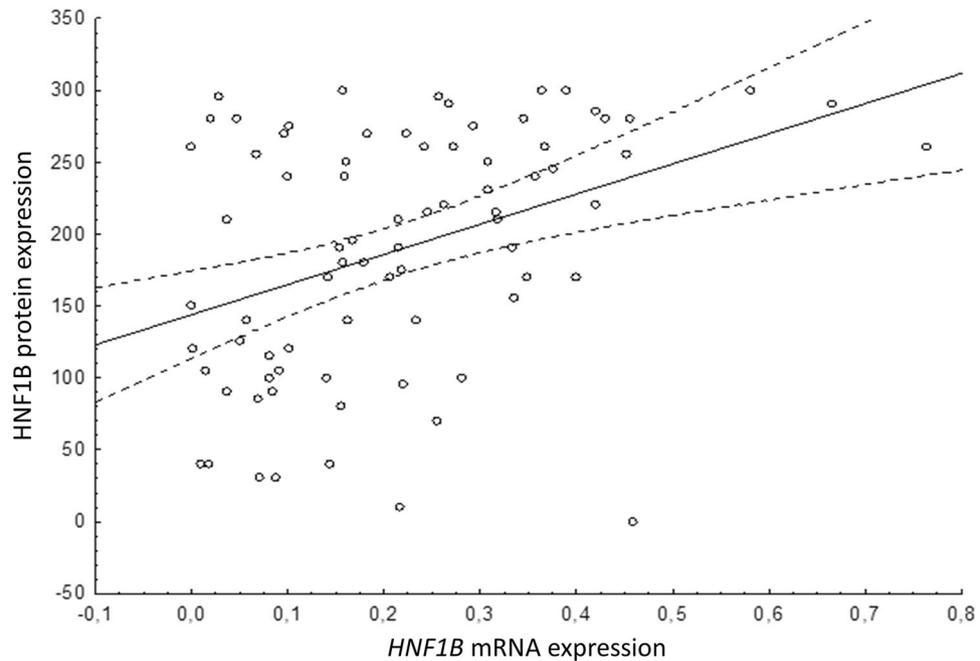
Moreover, we did not detect any major expression deviations of analysed ASVs in our dataset which could potentially indicate aberrant splicing resulted from the variant located in these sites. Although we cannot exclude the potential splicing effect of deep intronic variants, our findings suggest that all analysed splicing variants were created based on the alternative splicing mechanism and described differences in ASVs expression between T and NT sample sets are most probably resulting from the changes in alternative splicing.

#### Capture RNA-Seq confirmed the *HNF1B* quantitative and qualitative splicing pattern determined by ddPCR.

The RNA from six selected paired tissue samples with detected differences in the splicing patterns between T and NT were sequenced by the capture RNA-Seq approach. A representative example of one kidney T sample is visualised in Fig. 5. In all 12 representative samples reads were identified which corresponded to 3p,  $\Delta 7$ ,  $\Delta 7-8$ , and  $\Delta 8$  ASVs. Normalized NGS data (ratio of individual splicing reads) showed a similar relative expression of the studied variants when compared to ddPCR quantitative analysis (Table 3). Moreover, we did not identify any novel *HNF1B* splicing variants by capture RNA-Seq other than those reported in our previously published work, where deep sequencing of individual *HNF1B* exon-exon junctions was used, which could potentially miss the ASVs<sup>19</sup>.

#### TCGA SpliceSeq database data show similar expression of analysed *HNF1B* ASVs.

For further validation of our results, we used RNA-Seq based data from TCGA SpliceSeq database<sup>20</sup> (TCGAss) and compared them with our ddPCR expression results. Although the TCGAss data are limited due to the whole transcriptome sequencing with rather lower sequence depth which was not designed for thorough quantitative analysis of low expressed alternative transcripts in genes such as *HNF1B*, we were able to directly assess the expression of variants 3p,  $\Delta 7-8$  and  $\Delta 8$  in paired T and NT datasets that correspond to ours (Table 4). Variant  $\Delta 7$  as a simple exon skipping variant was absent in the downloadable TCGAss data probably due to the low coverage/number of supporting reads.



**Figure 3.** Correlation of the overall *HNF1B* mRNA and protein expression in 78 large intestine carcinoma samples showed weak positive correlation ( $R=0.39$ ;  $p<0.001$ ). *HNF1B* mRNA expression (X axis) is relative to *POLR2A* (*POLR2A* expression = 1). *HNF1B* protein expression (Y axis) is calculated as H-score. Each dot represents one sample. Dashed lines indicate the 95% confidence intervals for the regression line.

Comparison shows similar expression levels, usually in units of percent, in most cases with several exceptions. Most distinctive exception was  $\Delta 7-8$  expression in kidney NT tissue with median of 14.9% (compared to 1% measured by ddPCR) and 3p expression in prostate NT tissue with median 18.2% (compared to 29.1% measured by ddPCR).

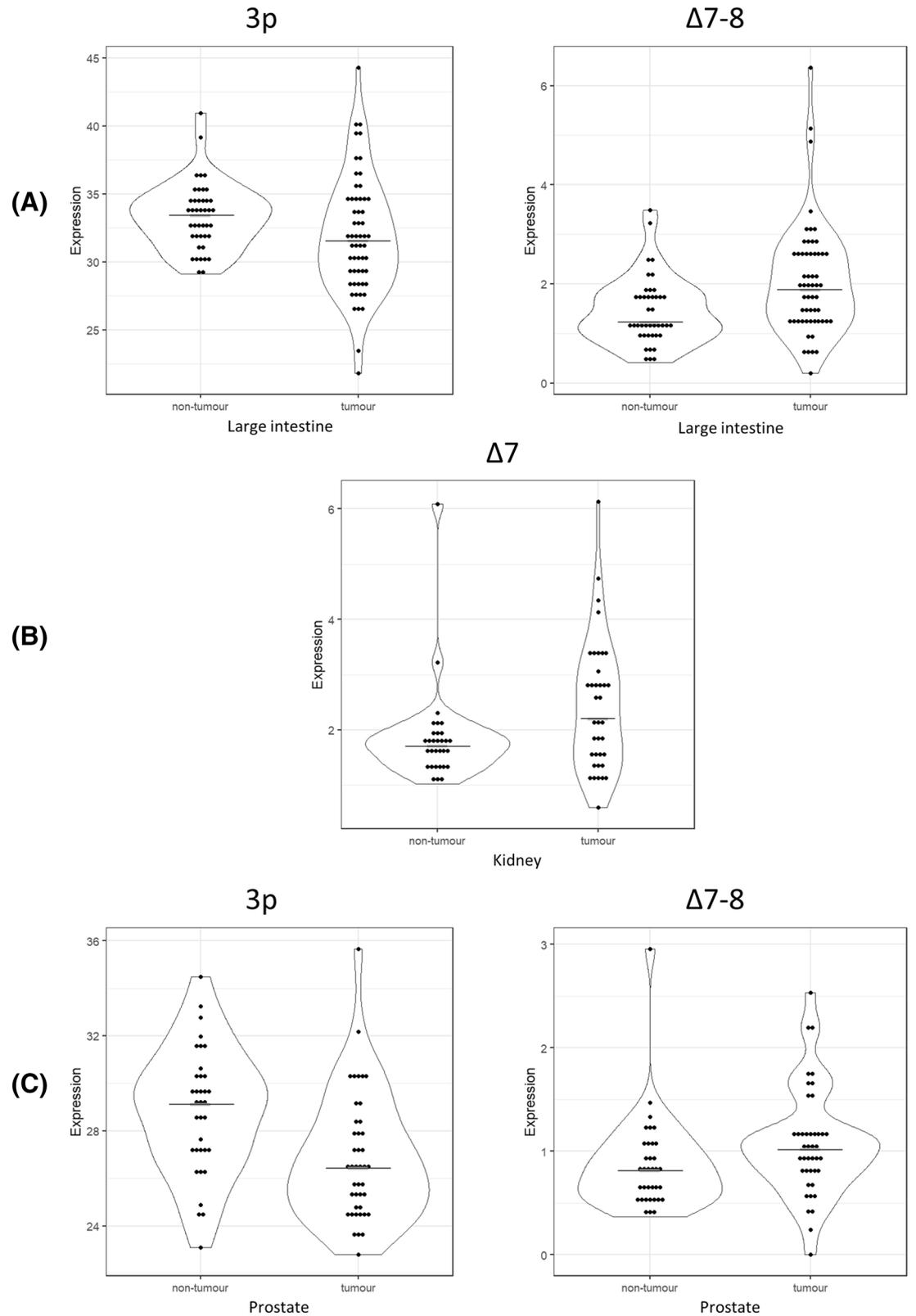
## Discussion

Although the significance of *HNF1B* in tumorigenesis has recently been closely analysed and discussed, its exact role and mechanism of action have not yet been fully clarified. There are studies suggesting that *HNF1B* may act as an oncogene in specific cancer types, such as ovarian clear cell carcinoma<sup>5,6,21</sup> and papillary renal cell carcinoma<sup>16</sup>. Other studies suggested that *HNF1B* acts mainly as a tumour suppressor in colorectal, prostate, ovarian, and some other types of solid tumours<sup>3,15,17,18,22</sup>. One of the possible explanations for this ambivalent character is the existence of one or more alternative splicing variants with dysregulated expression in tumour tissues which have either a regulatory role or code for *HNF1B* protein isoforms with different function<sup>19</sup>.

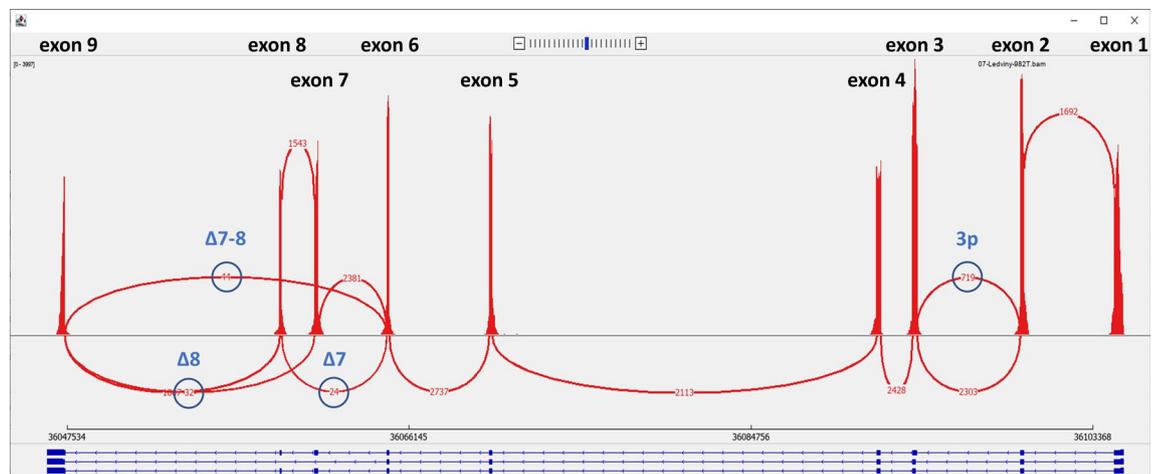
Our previous work showed that the splicing mechanism in ovarian, colorectal, kidney, pancreatic, and prostate tissue types can create a wide range of *HNF1B* alternative splicing events. Out of the 45 previously described ASVs of *HNF1B* mRNA, variants 3p,  $\Delta 7$ ,  $\Delta 7-8$ , and  $\Delta 8$  occurred in all analysed tissue pools in a higher portion, and variants  $\Delta 5-8$  and  $\Delta 6-8$  occurred in a majority of the tested tissue pools in moderate portion<sup>19</sup>. Moreover, according to our previous findings and the RefSeq database information (accessed April 30, 2021), all these ASVs have fully defined open reading frame and thus they potentially code for protein products (Fig. 1)<sup>19</sup>. Therefore, these ASVs were selected for further investigation by a precise ddPCR quantitative approach on a large-scale tumour and non-tumour sample set presented in this study. We confirmed the previous assumption that variants 3p,  $\Delta 7$ ,  $\Delta 7-8$ , and  $\Delta 8$  are expressed ubiquitously. These variants were detected in all 146 analysed NT tissue samples. The variant 3p corresponded to approximately 30% of all *HNF1B* transcripts, ASV  $\Delta 8$  was detected in 2–7% of all *HNF1B* transcripts, and variants  $\Delta 7$  and  $\Delta 7-8$  were detected in relatively low levels in approximately 1–2% of all *HNF1B* transcripts (Table 1). On the other hand, the previously described variants  $\Delta 5-8$  and  $\Delta 6-8$  were identified only in a portion of the tested samples and showed expression levels lower than 0.5% of the overall *HNF1B* mRNA expression, which is in accordance with our previous ASVs overview<sup>19</sup> and emphasizes the rather lower significance of these *HNF1B* ASVs.

Moreover, the TCGAss database information<sup>20</sup>, which is based on the TCGA RNA-Seq data (<https://bioinformatics.mdanderson.org/TCGASpliceSeq>), supports the presence of all four described variants in the majority of other tumour tissues if the expression filter for minor splice variants in the database (Min minor splice expression) is decreased from default 10% to 1%.

Further analysis of the alternative *HNF1B* transcripts revealed that the ASVs  $\Delta 7$ ,  $\Delta 7-8$ , and  $\Delta 8$  occur in alternative transcripts separately or in combination with ASV 3p. This information completes the look at the *HNF1B* splicing pattern and supports the existence of predicted *HNF1B* transcripts containing full-length exon 3 and  $\Delta 7$  variant (XM\_01525161.1); full-length exon 3 and  $\Delta 8$  variants (XM\_011525164.1); alternative 3p and  $\Delta 8$



**Figure 4.** Significantly different expression levels of the *HNF1B* alternative splicing variants in T and NT tissue samples. **(A)** Large intestine (NT = 42 samples; T = 57 samples); **(B)** kidney (NT = 31 samples; T = 37 samples); **(C)** prostate (NT = 35 samples; T = 42 samples). Data is visualized as violin plots. Each dot represents one sample. Expression is relative to overall *HNF1B* mRNA expression (100). Black line represents median. Complete results are visualized in Fig. S2.



**Figure 5.** Representative example of the *HNF1B* splicing pattern based on the capture RNA-Seq of a kidney T sample visualised as a sashimi plot in IGV (Broad Institute). The height of the peaks shows exon coverage, the red lines show splicing reads, the red numbers in the red lines show the number of splicing reads, the ASVs are indicated in blue font and the number of ASVs' reads is circled. The dark blue scheme under the plot shows the gene structure of the protein coding *HNF1B* transcripts (blue boxes indicate exons; thin horizontal lines indicate introns). The cut-off for visualisation of the splicing reads was set to 10. The results are provided in Table 3.

	3p		Δ7		Δ7–8		Δ8		ddPCR	NGS
	ddPCR (%)	NGS (%)	tmp/μl	cov						
Pancreas T	29.9	27.5	2.8	1.5	2.6	3.2	10.4	5.7	385	3766
Pancreas NT	30.2	30.5	3.9	1.7	2.8	4.2	11.4	5.2	165	1409
Large intestine T	33.2	32.8	3.9	3.7	5.1	5.6	15.1	6.9	214	2572
Large intestine NT	32.7	32.5	1.8	0.7	1.2	1.2	6.0	3.9	281	2775
Prostate T	24.4	21.2	4.4	1.8	2.2	2.5	4.8	2.5	66	663
Prostate NT	29.7	27.2	1.2	1.8	0.5	0.4	2.7	1.0	158	2490
Kidney T	30.1	23.8	4.7	0.8	2.7	1.5	6.9	1.1	916	3897
Kidney NT	28.4	25.2	1.2	1.1	0.6	1.0	2.3	1.9	981	11,619
Endometrium T	27.4	25.9	0.4	0.1	0.8	0.7	1.5	0.3	170	2301
Endometrium NT	31.5	24.4	1.6	0.8	0.8	0.9	2.7	2.3	203	1786
Ovary T	19.8	16.4	1.8	2.4	2.1	2.0	9.1	3.3	137	866
Ovary NT	30.0	24.3	2.0	1.2	1.1	2.0	5.2	3.5	113	1082

**Table 3.** The proportion of alternative splicing variants calculated from ddPCR-based and NGS-based data in selected paired tumour and non-tumour tissues (with raw data overview). Numbers represent the percentage of each individual variant related to the overall *HNF1B* mRNA expression (calculated as the sum of canonical exon 3 and 3p variants in both approaches). Each line represents data from one representative sample; ddPCR tmp/μl represents number of *HNF1B* templates per 1 μl of cDNA detected by ddPCR; NGS cov. represents peak coverage of *HNF1B* exon 3. T tumour tissue, NT non-tumour tissue.

variants (XM\_011525160.1), and further suggests the existence of full-length exon 3 and Δ7–8 variants (novel) and alternative exon 3p and Δ7 variants (novel).

Other predicted *HNF1B* transcripts XM\_011525162.2 and XM\_011525163.2 (which contain canonical exons 1–4 and are concluded by the exonization of a part of intron 4) were not detected in any samples analysed by the NGS approach. This suggests that these transcripts are either expressed in low levels (<1%) and thus under the detection limit of the used NGS method, or specific for a type of tissue which was not analysed in our study.

Previous *HNF1B* DNA mutation analysis of presented sample set<sup>15–17</sup> did not reveal any variant located in canonical or potential cryptic splice sites nor the variant with potential splice effect in coding sequence and exon flanking areas of the *HNF1B* gene. Even though potential variants with splice effect located deep in the introns were not analysed, based on even expression pattern of ASVs, we can conclude, that detected frequencies of analysed *HNF1B* ASVs are based on the alternative splicing rather than aberrant splicing mechanism.

So far, only two studies have analysed the expression of *HNF1B* ASVs. Specifically, they compared only the canonical variant [i.e., full-length *HNF1B* transcript, named by the authors as *HNF1B*(A)] and the alternative transcript 3p variant [named by the authors as *HNF1B*(B)]<sup>23,24</sup>. Our results are not in full accordance with those authors. When they compared the 3p variant to the full-length *HNF1B* mRNA (by qPCR normalization to reference B2M mRNA expression), they observed approximately 60% lower levels of the 3p ASV in the

	3p		Δ7-8		Δ8		ddPCR	TCGAss
	ddPCR	TCGAss	ddPCR	TCGAss	ddPCR	TCGAss	N	N
Large intestine T	31.6%	24.5%	1.9%	1.0%	5.8%	4.3%	33	33
Large intestine NT	33.5%	27.0%	1.2%	0.7%	6.7%	3.1%		
Kidney T	29.2%	24.0%	1.4%	0.9%	4.0%	2.4%	28	57
Kidney NT	28.2%	32.5%	1.0%	14.9%	3.8%	2.0%		
Prostate T	26.5%	20.7%	1.0%	0.0%	2.6%	0.0%	28	38
Prostate NT	29.1%	18.2%	0.8%	0.0%	2.3%	2.1%		

Numbers represent the percentage of each individual variant. ddPCR based expression is related to the overall *HNF1B* mRNA expression (calculated as the sum of canonical exon 3 and 3p variants), TCGAss expression is calculated as a 100-PSI / percent splice in; see methods). Each line represents data from corresponding T and NT datasets. T – tumour tissue, NT – non-tumour tissue; N – number of samples in datasets.

**Table 4.** Comparison of *HNF1B* ASVs expression measured by our ddPCR approach with TCGAss. Numbers represent the percentage of each individual variant. ddPCR based expression is related to the overall *HNF1B* mRNA expression (calculated as the sum of canonical exon 3 and 3p variants), TCGAss expression is calculated as a 100-PSI/percent splice in; see methods). Each line represents data from corresponding T and NT datasets. T tumour tissue, NT non-tumour tissue, N number of samples in datasets.

pancreatic islets, a balanced expression in the kidney tissues, and approximately 50% higher expression in the pancreas<sup>23</sup>. Based on our ddPCR approach we found that the median expression of ASV 3p is about 50% lower than the expression of the full-length exon 3 variant in all 142 analysed NT tissues (Table 1). In their later study, Harries et al. showed that the expression of the full-length *HNF1B* transcript is non-significantly slightly lower (approximately 0.8×), and the expression of the 3p variant is significantly higher (approximately 7×) in 21 prostate adenocarcinoma tissue samples when compared with 39 non-malignant (benign hyperplasia) tissues (by qPCR normalization to B2M and GUSB reference mRNA transcripts)<sup>24</sup>. Based on our current study, we can state that the overall *HNF1B* expression is similarly lower (approximately 0.9×) in prostate T samples compared with NT samples with the marginal statistical significance (Fig. 2C, Table 2A). In contrast to the results published by Harries et al., our data showed that the expression of the 3p variant is significantly lower in prostate adenocarcinomas when compared with prostate NT samples (Fig. 4C, Table 2B). However, it cannot be reliably macroscopically distinguished between benign hyperplasia and non-benign character of samples in our NT prostate sample group even though all tissue specimens were evaluated by the trained pathologists. Although we do not expect major differences of the *HNF1B* ASVs expression in healthy prostate tissue samples compared to benign hyperplasia prostate tissue samples and our data from individual NT samples did not deviate, our results may be affected by this issue.

To our knowledge, neither *HNF1B* overall expression in other tissue types nor the expression of individual alternative transcripts has been previously analysed or published by other authors. However, we were able to compare our data to TCGAss database (Table 4) which provides transcript splicing patterns on the base of The Cancer Genome Atlas project (TCGA) RNA-Seq datasets. Direct comparison of the *HNF1B* ASVs expression results showed certain level of similarity, despite the different approaches for sample analysis and ASVs frequency calculation.

Unfortunately, TCGA RNA-Seq sequencing depth was not sufficient for the unbiased statistical evaluation of the *HNF1B* ASVs expression in T versus NT tissue samples and its comparison to our results obtained by sensitive ddPCR results. The TCGAss data showed approx. 4× lower coverage depth in exon-exon junctions in comparison with our panel RNA-Seq (Table 3). We consider this as insufficient for the reliable low-expressed ASVs detection which is reflected at Table 4. Expression of variant Δ7 was not evaluated by TCGAss in individual samples (see methods), thus it is not reported there and variants Δ7–8 and Δ8 showed minimal expression in prostate T and NT sets, where is the overall *HNF1B* expression lowest, which is most likely caused by insufficient sensitivity of used method. These limitations could also be the reason for the observed differences in Table 4.

Our data revealed that the only tissue with a strong significant difference of the overall *HNF1B* expression when comparing NT with T samples was the large intestine (Fig. 2A). Our analysis showed that the significantly lower overall *HNF1B* mRNA expression in the large intestine carcinoma samples correlates with low protein expression in the same sample set (Fig. 3)<sup>15</sup>. Given the previous results gained from the same set of large intestine carcinoma samples which showed that the effect of the *HNF1B* gene promoter methylation and the *HNF1B* mutation is minor<sup>15</sup>, we can propose that the decreased *HNF1B* mRNA expression in the large intestine carcinoma samples is most probably caused by a reduction of the *HNF1B* transcriptional rate, which is usually directly affected by upstream gene expression regulators.

Another important mechanism of the potential reduction of the overall *HNF1B* protein function could be the effect of alternative splicing. A significant increase in the expression of the Δ7–8 variant was observed in the large intestine T samples when compared with the NT samples, by both statistical approaches (Fig. 4A,C; Table 2B). In the kidney T samples, a significant increase of the Δ7 variant expression was detected compared with the kidney NT samples (Fig. 4B; Table 2B). The transcript Δ7–8, as well as the Δ7 or Δ8 variants, lacks the part which codes an important domain responsible for the transcriptional activation of *HNF1B* gene targets.

The same DNA binding capacity but lower transactivation activity of the proteins raised from the *HNF1B*  $\Delta 7$ – $8$ / $\Delta 7/\Delta 8$  transcripts could be expected, but a functional analysis of these predicted proteins is needed prior to the final evaluation of their role and effect.

Interestingly, the variant 3p, which lacks 78 bp at its 5' end of exon 3 (which codes for a spacer between the first and the second DNA binding domain of the HNF1B protein) shows a significantly lower expression in the large intestine T samples when compared with the NT samples (Fig. 4A). A similar observation was made in the prostate T samples, where a significantly lower expression was also detected compared with the NT samples (Fig. 4C; Table 2B). Regarding the high expression of the 3p ASV of all *HNF1B* transcripts, we can assume that this transcription variant codes for a protein with a similar transcriptional activation potential. However, as with other variants, a functional assay should be done prior to the final evaluation of the HNF1B 3p isoform protein function and its role in tumorigenesis.

Although we described statistically significant differences in the expression of certain ASVs in the T samples compared with the NT samples, the variances are relatively low in summary. Therefore, changes in the *HNF1B* splicing pattern are probably not the key mechanism in the regulation of gene expression of fully functional HNF1B protein in the analysed tumour tissues. Observed changes in the *HNF1B* splicing pattern may be a consequence of the overall tumour splicing disbalance rather than presence of potential deep intronic variants with splicing effect. However, this assumption should be confirmed by a suitable large-scale method such as RNA-Seq, where the multiple splicing patterns could be targeted and evaluated at once.

Finally, another interesting part of this work was the direct comparison of the *HNF1B* splicing data acquired by both ddPCR and capture RNA-Seq (Table 3). This comparison showed that a well-chosen normalization (in this case the sum of the canonical exon 3 and alternative exon 3p reads in NGS and expression in ddPCR) provides similar quantitative results for the variants 3p,  $\Delta 7$ , and  $\Delta 7$ – $8$  in a majority of the analysed tissue types. Quantification of the  $\Delta 8$  variant showed greater variability across tissues, which was probably caused by non-ideal probes binding or NGS reads mapping in this area. Nevertheless, we believe that additional fine-tuning of the sample preparation and NGS data analysis can lead to a true quantitative analysis of the splicing pattern by the RNA capture NGS approach. Further, capture-based results in contrast with the whole transcriptome data usually benefits from higher coverage of low-expressed targeted genes, such as *HNF1B*. In comparison with external databases such as TCGA or GTEx Portal<sup>7</sup>, we showed approx. 4 × greater coverage of the *HNF1B* exon-exon junctions which allows us to precisely calculate the expression even for the minor (< 10%) alternative transcripts. In addition to the other information gained by RNA NGS approach, analysis of the splicing pattern in a variety of genes can be beneficial. The capture RNA-Seq approach would bring increased throughput and decreased economic demands compared to the otherwise used large-scale quantitative splicing experiments.

In conclusion, our work unravels the spectrum of *HNF1B* alternative splicing. We described the spectrum of *HNF1B* ASVs in non-malignant tissues and the quantitative changes of the *HNF1B* ASVs in corresponding tumours. The exact function of individual ASVs is not fully clarified and should be thoroughly functionally determined. Our work provided a guide for the prioritization of alternative variants/isoforms for such experiments. Moreover, we described quantitative changes on the level of overall *HNF1B* mRNA expression and showed a positive correlation between decreased mRNA level and decreased protein level in colorectal cancer. Additionally, we showed that the NGS approach of capture RNA-Seq could be beneficial for the collective evaluation and direct quantification of the analysed splicing patterns.

## Materials and methods

**Patients and samples.** Samples for the study were provided by The Bank of Biological Material, First Faculty of Medicine, Charles University. The tissue samples were collected during surgical procedures and processed by trained pathologists, including macroscopic evaluation of the whole resected tissue specimen prior to fixation. Representative tumour and paired non-tumour tissue samples (taken from the periphery of each resected specimen where available) were stored in RNAlater (Thermo Fisher) according to the manufacturer's instructions until the genetic material had been isolated. The total amount of 512 tumour (T) and non-tumour (NT) tissue samples were obtained from the pancreas (14 adenocarcinomas; 9 NT), large intestine (83 adenocarcinomas; 82 NT), prostate (73 adenocarcinoma; 49 NT), kidney (59 clear cell renal cell adenocarcinomas; 43 NT), and female genital tract (100 NT). Non-tumour female genital tract samples were collected from the ovary, fallopian tubes, endometrium, and cervix; therefore, they are collectively referred to as "Female internal genital tract NT samples". One high-grade serous carcinoma sample and one endometrioid endometrial carcinoma sample were additionally included into the dataset due to the evaluation of NGS approach.

The study has been approved by The Ethics Committee of General University Hospital in Prague in compliance with the Helsinki Declaration (ethical approval number 41/16 as a part of the grant from the Czech Research Council 17-28404A) and all experiments were performed in accordance with these guidelines and regulations. The ethics committee which approved this retrospective study waived the need for informed consent.

**Total RNA isolation, quality control, and cDNA synthesis.** The samples were processed according to the Digital MIQE Guidelines<sup>25</sup> and as described before<sup>17,19</sup>. Briefly, the total RNAs and DNAs were isolated from the homogenized part of thawed tissue (10–30 mg) according to the Simultaneous Purification of Genomic DNA and Total RNA from Animal Tissues protocol by using an AllPrep DNA/RNA Mini kit (Qiagen).

All isolated RNA samples were quantified by the NanoDrop 2000 instrument (Thermo Fisher) and the RNA integrity was characterized by an RNA Quality Number (RQN) using the Fragment Analyzer capillary electrophoresis system (AATI) and Standard RNA kit (Agilent; tissue samples RQN = 9.2; range 5–10). Samples with RQN < 7.5 and c < 25 ng/μl of total RNA were excluded from further analysis.

Prior to the single strand cDNA synthesis, 3.75 µg of RNA of each sample was treated by DNase I (Thermo Fisher) according to the manufacturer's instructions. Reverse transcription was performed in a 40 µl reaction using SuperScript III Reverse Transcriptase (Thermo Fisher) with random hexamers (Roche) as described previously<sup>26</sup>. All cDNA samples were stored in -20 °C until the quantification step (performed within 1 month).

**Analysis of the overall *HNF1B* expression and *HNF1B* alternative splicing variants expression.** The overall *HNF1B* expression analysis of cDNAs was performed by droplet digital PCR (ddPCR) approach using QX200 ddPCR system (Bio-Rad; including Automated Droplet Generator instrument) and EvaGreen quantification kits (Bio-Rad) as described by Dundr et al.<sup>17</sup>.

Several thorough optimization steps were performed prior to the analyses. These included the testing of expression of three pre-selected reference gene mRNA targets (*POLR2A*, *HPRT1*, and *ATP5F1B*; Table S1) and two *HNF1B* mRNA targets (in the 5' UTR and 3' UTR; Table S1), confirming the repeatability, reproducibility, optimal primers annealing temperature, and specificity (by Sanger sequencing). The reactions were prepared using QX200 ddPCR EvaGreen Supermix (Bio-Rad; according to the manufacturer's instructions), 1 µl of cDNA template (which corresponded to approximately 90 ng of the total RNA), and 4 pmol of each of the primer pairs (200 nM final concentration) in 20 µl reaction volume.

The *HNF1B* ASVs expression analysis was performed by the same ddPCR system using probe kits as well as custom-designed FAM/HEX probes specific for unique and canonical exon-exon connections in one reaction for the direct evaluation of the alternative transcript amounts. Based on our previous research<sup>19</sup>, six major transcription variants of *HNF1B* were quantified (3p; Δ7; Δ7-8; Δ8; Δ5-8; and Δ6-8). For each variant, a pair of primers and probes were designed (Table S1). For example, the reaction which represents the Δ7 variant included the forward primer in exon 6 and the reverse primer in exon 8, the FAM probe specific for alternative exon 6 and exon 8 junction, and the HEX probe specific for canonical exon 6 and exon 7 junction. The optimization steps were performed similarly to the process described above, with the addition of a probe multiplexing step for each reaction. The reactions were prepared using ddPCR Supermix for Probes (No dUTP; Bio-Rad; according to the manufacturer's instructions), 1 µl of cDNA template (approx. 90 ng of total RNA), 5 pmol of each of the two primers and two probes (250 nM final concentration) in a 20 µl reaction volume. Droplets were generated in the QX200 AutoDG instrument (Bio-Rad) according to the general instructions and amplified according to the manufacturer's protocol (10 min incubation at 95 °C followed by 40 cycles of 94 °C for 20 s, 58 °C for 45 min, and final 98 °C for 10 min).

Data was acquired by the QX200 Droplet Reader instrument (Bio-Rad), using the standard acquisition protocol for Eva-Green/Probe master mix and analysed by the QuantaSoft software (Bio-Rad). The threshold for positive droplet signals of each of the individual amplicons was set as the average of the thresholds which were calculated automatically by QuantaSoft software during the optimization steps (48 reactions for each amplicon). The thresholds of all the acquired targets were manually confirmed. The final data of targets, expressed as the number of templates in 20 µl of master mix (which corresponded to 1 µl of cDNA), was recalculated in the following manner: (i) as the number of targets per reference *POLR2A* (overall *HNF1B* mRNA expression); (ii) as the percentage of the overall *HNF1B* variants (ASV relative expression; overall *HNF1B* expression was calculated as the sum of the full-length exon 3 variant and the 3p variant). The overall *HNF1B* expression was determined as the sum of the full-length exon 3 variant and the alternatively spliced 3p variant, which correlates to the overall *HNF1B* expression determined as the number of templates measured by the amplicon in 5' UTR. The final recalculated data was analysed as described in the Statistics section.

The samples with *POLR2A* expression < 50 templates in 1 µl of cDNA were excluded from further statistical analysis of the overall *HNF1B* expression. The mean of the *POLR2A* expression in the whole dataset was 999 templates per 1 µl of cDNA (min = 58; max = 2076).

The samples with combined alternative exon 3 and canonical exon 3 expression > 20 templates/1 µl cDNA were evaluated as samples with sufficient *HNF1B* expression for further statistical analysis of individual ASVs expression. The mean of the *HNF1B* combined exon 3 and 3p expression in the subset was 141 templates per 1 µl of cDNA (min = 23; max = 3868).

**Analysis of ASV's combination.** Complementary DNA (cDNA) for the analysis of the potential combination of splicing events was synthesized using eight non-tumour RNA samples from the kidney as a template. The samples were equimolarly pooled into a mixture of 4 µg of total RNA, which was treated by DNase I, prior to cDNA synthesis as described in Hojny et al.<sup>19</sup>. The reverse transcription was performed according to the manufacturer's protocol of SuperScript III Reverse Transcriptase (Thermo Fisher) in 40 µl reaction volume by using 1 µl of oligo dT primer (final concentration 1.25 µM). The created cDNA was amplified using forward primer in exon 2 and one of three reverse primers, which were specific for the unique exon-exon junctions of each of Δ7, Δ7-8, and Δ8 variant (the final concentration of primers was 200 nM; Table S1). The PCR reactions were mixed into 10 µl by using 5 × HOT FIREPol EvaGreen HRM Mix and amplified according to the instructions (95 °C—12 min; 35 cycles of 95 °C—15 s, 62 °C—20 s, 72 °C—1 min). The samples were visualized by 1% agarose gel electrophoresis. The PCR products were sequenced and confirmed by the direct Sanger sequencing.

**Statistical analyses.** All statistical analyses were performed using the software Statistica v.13.5.0 (TIBCO Statistica) and/or the R software v.4.0.2<sup>27</sup>. The package “ggplot2”, implemented in R software, was used for data visualisation (violin plots, Figs. 2, 4). The Shapiro–Wilk test was used to control data normality. The non-parametric ANOVA approach (Mann–Whitney *U* test) was used to evaluate the associations between the *HNF1B* expression (overall or ASVs) and the type of tissue (tumour vs non-tumour). For the pairwise comparisons (matched T and NT samples), Wilcoxon matched-pairs signed-ranks test was performed. The *HNF1B* protein

expression (H-score) and RNA expression were correlated using Pearson's method. All tests were two-sided and a p-value of less than 0.05 was considered as significant. All summary values were expressed as median.

The sample set from the pancreas was excluded from the statistical evaluation due to the insufficient number of samples suitable for analysis (10 T; 8 NT).

**Capture RNA-Seq library preparation, sequencing, and data analysis.** The representative paired T and NT samples of RNA from six different tissues (pancreas, large intestine, prostate, kidney, EEC with paired tissue samples, and HGSC with paired tissue samples) were prepared similarly as described by Walker et al.<sup>28</sup> using the standard KAPA RNA HyperPrep kit, enriched by the SeqCap protocol (Roche) by using a 257 kbp panel of gene targets (NimbleGen, Roche)<sup>17</sup>. The data was analysed by CLC Genomics Workbench v.21.0.3. software (Qiagen) by a standard RNA analysis pipeline. The mapped data was manually analysed in IGV (Broad Institute)<sup>29</sup>, and the identified splicing reads (reads which started in one exon and ended in another) were recalculated in the same way as the data from ddPCR—as a percentage of the overall *HNF1B* expression (overall *HNF1B* expression was calculated as the sum of splicing reads of the full-length exon 3 variant and the 3p variant).

**TCGA SpliceSeq data.** The TCGA SpliceSeq<sup>20</sup> (<https://bioinformatics.mdanderson.org/TCGASpliceSeq>) *HNF1B* data (based on whole transcriptome RNA-Seq approach) from colon adenocarcinoma (COAD), kidney renal clear cell carcinoma (KIRC) and prostate adenocarcinoma (PRAD) datasets together with their paired NT tissue samples was downloaded directly from TCGAss download page. Non-paired T samples and samples without detected ASV 3p expression (which is considered ubiquitous) were filtered out. The PSI values (percent splice in; percent of transcripts where the exon or variant is present) were recalculated as 100% – PSI which corresponds to a percentage where the variant is missing. Altogether paired T and NT samples of large intestine (COAD; N = 33); kidney (KIRC; N = 57) and prostate (PRAD; N = 38) datasets were further analysed. Downloaded data lacked variant  $\Delta 7$ , as a single exon skipping between exons 6 and 8. This variant was detected only in fraction of TCGA samples as the web viewer showed. In TCGAss, another variant 7 is present, but its expression is related to the exon 6–9 splice junction as skipped exon 7 together with skipped exon 7–8 variant. This calculation is not compliant with our approach and therefore we skipped this variant for further comparison.

**Database information.** The information about the *HNF1B* transcripts was obtained from the RefSeq database as of April 30, 2021, where the spectrum of the fully characterized and predicted *HNF1B* transcripts was wider in comparison to the Ensembl database. The general information about the *HNF1B* mRNA expression in a variety of human tissues, as well as other used reference genes for mRNA expression, was obtained from the GTEX Portal database<sup>7</sup>.

### Data availability

The data generated or analysed in this current study is included in this article (and its Supplementary Information Files). The raw data supporting the conclusions of this manuscript will be available from the corresponding author upon reasonable request to any qualified researcher.

Received: 19 May 2021; Accepted: 13 December 2021

Published online: 07 January 2022

### References

- Cereghini, S. Liver-enriched transcription factors and hepatocyte differentiation. *FASEB J.* **10**, 267–282 (1996).
- Barbacci, E. et al. HNF1beta/TCF2 mutations impair transactivation potential through altered co-regulator recruitment. *Hum. Mol. Genet.* **13**, 3139–3149. <https://doi.org/10.1093/hmg/ddh338> (2004).
- Bartu, M. et al. The role of HNF1B in tumorigenesis of solid tumours: A review of current knowledge. *Folia Biol. (Praha)* **64**, 71–83 (2018).
- Yu, D. D., Guo, S. W., Jing, Y. Y., Dong, Y. L. & Wei, L. X. A review on hepatocyte nuclear factor-1beta and tumor. *Cell Biosci.* **5**, 58. <https://doi.org/10.1186/s13578-015-0049-3> (2015).
- Suzuki, E. et al. Transcriptional upregulation of HNF-1beta by NF-kappaB in ovarian clear cell carcinoma modulates susceptibility to apoptosis through alteration in bcl-2 expression. *Lab Invest.* **95**, 962–972. <https://doi.org/10.1038/labinvest.2015.73> (2015).
- Tsuchiya, A. et al. Expression profiling in ovarian clear cell carcinoma: Identification of hepatocyte nuclear factor-1 beta as a molecular marker and a possible molecular target for therapy of ovarian clear cell carcinoma. *Am. J. Pathol.* **163**, 2503–2512 (2003).
- Carithers, L. J. et al. A novel approach to high-quality postmortem tissue procurement: The GTEx project. *Biopreserv. Biobank.* **13**, 311–319. <https://doi.org/10.1089/bio.2015.0032> (2015).
- Debiais-Delpech, C. et al. Expression patterns of candidate susceptibility genes HNF1beta and CtBP2 in prostate cancer: Association with tumor progression. *Urol. Oncol.* **32**, 426–432. <https://doi.org/10.1016/j.urolonc.2013.09.006> (2014).
- Elliott, K. S. et al. Evaluation of association of HNF1B variants with diverse cancers: Collaborative analysis of data from 19 genome-wide association studies. *PLoS One* **5**, e10858. <https://doi.org/10.1371/journal.pone.0010858> (2010).
- Glinsky, G. V., Glinskii, A. B., Stephenson, A. J., Hoffman, R. M. & Gerald, W. L. Gene expression profiling predicts clinical outcome of prostate cancer. *J. Clin. Invest.* **113**, 913–923. <https://doi.org/10.1172/JCI20032> (2004).
- Janky, R. et al. Prognostic relevance of molecular subtypes and master regulators in pancreatic ductal adenocarcinoma. *BMC Cancer* **16**, 632. <https://doi.org/10.1186/s12885-016-2540-6> (2016).
- Kondratyeva, L. G., Chernov, I. P., Zinovyeva, M. V., Kopantzev, E. P. & Sverdlov, E. D. Expression of master regulatory genes of embryonic development in pancreatic tumors. *Dokl. Biochem. Biophys.* **475**, 250–252. <https://doi.org/10.1134/S1607672917040020> (2017).
- Spurdle, A. B. et al. Genome-wide association study identifies a common variant associated with risk of endometrial cancer. *Nat. Genet.* **43**, 451–454. <https://doi.org/10.1038/ng.812> (2011).

14. Yamamoto, S. *et al.* Immunohistochemical detection of hepatocyte nuclear factor 1beta in ovarian and endometrial clear-cell adenocarcinomas and nonneoplastic endometrium. *Hum. Pathol.* **38**, 1074–1080. <https://doi.org/10.1016/j.humpath.2006.12.018> (2007).
15. Bartu, M. *et al.* Expression, epigenetic, and genetic changes of HNF1B in colorectal lesions: An analysis of 145 cases. *Pathol. Oncol. Res.* **26**, 2337–2350. <https://doi.org/10.1007/s12253-020-00830-2> (2020).
16. Bartu, M. *et al.* Analysis of expression, epigenetic, and genetic changes of HNF1B in 130 kidney tumours. *Sci. Rep.* **10**, 17151. <https://doi.org/10.1038/s41598-020-74059-z> (2020).
17. Dunder, P. *et al.* HNF1B, EZH2 and ECI2 in prostate carcinoma. Molecular, immunohistochemical and clinico-pathological study. *Sci. Rep.* **10**, 14365. <https://doi.org/10.1038/s41598-020-71427-7> (2020).
18. Nemejcova, K. *et al.* A comprehensive analysis of the expression, epigenetic and genetic changes of HNF1B and ECI2 in 122 cases of high-grade serous ovarian carcinoma. *Oncol. Lett.* **21**, 185. <https://doi.org/10.3892/ol.2021.12446> (2021).
19. Hojny, J. *et al.* Identification of novel HNF1B mRNA splicing variants and their qualitative and semi-quantitative profile in selected healthy and tumour tissues. *Sci. Rep.* **10**, 6958. <https://doi.org/10.1038/s41598-020-63733-x> (2020).
20. Ryan, M. *et al.* TCGASpliceSeq a compendium of alternative mRNA splicing in cancer. *Nucleic Acids Res.* **44**, 1018–1022. <https://doi.org/10.1093/nar/gkv1288> (2016).
21. Nemejcova, K. *et al.* Expression, epigenetic and genetic changes of HNF1B in endometrial lesions. *Pathol. Oncol. Res.* **22**, 523–530. <https://doi.org/10.1007/s12253-015-0037-2> (2016).
22. Buchner, A. *et al.* Downregulation of HNF-1B in renal cell carcinoma is associated with tumor progression and poor prognosis. *Urology* **76**(507), e506–e511. <https://doi.org/10.1016/j.urology.2010.03.042> (2010).
23. Harries, L. W., Brown, J. E. & Gloyn, A. L. Species-specific differences in the expression of the HNF1A, HNF1B and HNF4A genes. *PLoS One* **4**, e7855. <https://doi.org/10.1371/journal.pone.0007855> (2009).
24. Harries, L. W., Perry, J. R., McCullagh, P. & Crundwell, M. Alterations in LMTK2, MSMB and HNF1B gene expression are associated with the development of prostate cancer. *BMC Cancer* **10**, 315. <https://doi.org/10.1186/1471-2407-10-315> (2010).
25. Huggett, J. F. *et al.* The digital MIQE guidelines: Minimum information for publication of quantitative digital PCR experiments. *Clin. Chem.* **59**, 892–902. <https://doi.org/10.1373/clinchem.2013.206375> (2013).
26. Hojny, J. *et al.* Multiplex PCR and NGS-based identification of mRNA splicing variants: Analysis of BRCA1 splicing pattern as a model. *Gene* **637**, 41–49. <https://doi.org/10.1016/j.gene.2017.09.025> (2017).
27. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2020). <https://www.R-project.org>
28. Walker, L. C. *et al.* Comprehensive assessment of BARD1 messenger ribonucleic acid splicing with implications for variant classification. *Front. Genet.* **10**, 1139. <https://doi.org/10.3389/fgene.2019.01139> (2019).
29. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26. <https://doi.org/10.1038/nbt.1754> (2011).

## Acknowledgements

This research was funded by Ministry of Health, Czech Republic (Conceptual development of research organization, General University Hospital in Prague [grant number 64165]; by Charles University (Project Progres Q28/LF1, UNCE [grant number 204065] and SVV [grant number 260516]), and by the European Regional Development Fund (CZ.02.1.01/0.0/0.0/18\_046/0015959; BBMRI\_CZ LM2018125). The authors wish to thank to Zachary H. K. Kendall, BA (Institute for History of Medicine and Foreign Languages, First Faculty of Medicine, Charles University in Prague) for the English language corrections; Viktor Stránecký, PhD for the NGS data analysis support.

## Author contributions

Conceptualization: J.H., P.D. and I.S.; Methodology: J.H., R.M., E.K., Q.H.B. and P.K.; Software: J.H. and R.M.; Validation: M.B., K.N., M.K. and I.S.; Formal Analysis: J.H., R.M., E.K. and Q.H.B.; Investigation: J.H., M.B.; Resources: M.K., and P.D.; Data Curation: J.H., R.M., K.N. and I.S.; Writing—Original Draft Preparation: J.H. and R.M.; Writing—Review and Editing: M.B., P.K. and I.S.; Visualization, J.H. and R.M.; Supervision, I.S. and P.D.; Project Administration: J.H. and I.S.; Funding Acquisition, M.K. and P.D.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-03989-z>.

**Correspondence** and requests for materials should be addressed to I.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

# Metabolismo do glicogênio muscular durante o exercício físico: mecanismos de regulação

## *Muscle glycogen metabolism during exercise: mechanism of regulation*

Adriano Eduardo LIMA-SILVA<sup>1,2</sup>

Tony Charles FERNANDES<sup>2</sup>

Fernando Roberto DE-OLIVEIRA<sup>2</sup>

Fábio Yuzo NAKAMURA<sup>3</sup>

Monique da Silva GEVAERD<sup>2</sup>

### RESUMO

---

Uma série de estudos tem sido realizada para compreensão do metabolismo de glicogênio muscular durante o exercício. Estudos clássicos apontaram uma associação entre as reservas iniciais de glicogênio muscular e o tempo de sustentação do esforço. O glicogênio muscular diminui de forma semi-logarítmica em função do tempo, mas a concentração desse substrato não chega a zero, o que sugere a participação de outros mecanismos de fadiga na interrupção do exercício prolongado. Nesse tipo de atividade, a depleção de glicogênio, primeiro, ocorre nas fibras de contração lenta, seguida pela depleção nas de contração rápida. A diminuição na taxa de utilização de glicogênio muscular está sincronicamente ligada ao aumento no metabolismo de gordura, mas o mecanismo fisiológico é pouco compreendido. Estudos recentes sugerem que uma diminuição da insulina durante o exercício limitaria o transporte de glicose pela membrana plasmática, causando um aumento no consumo de ácidos graxos. Alguns estudos têm demonstrado, também, que a própria estrutura do glicogênio muscular pode controlar a entrada de ácidos graxos livres na célula, via proteína quinase. Fisicamente, a molécula de glicogênio se apresenta de duas formas, uma com estrutura molecular menor (aproximadamente,  $4,10^5$  Da, Proglucogênio) e outra maior (aproximadamente,  $10^7$  Da, Macroglucogênio). Aparentemente, a forma Proglucogênio é metabolicamente mais ativa no exercício e a Macroglucogênio mais suscetível a aumentar com dietas de supercompensação. Maior concentração de hipoxantinas e amônia no exercício com depleção de glicogênio muscular também foi relatada, mas estudos com melhor controle da intensidade do esforço podem ajudar a elucidar essa questão.

**Termos de Indexação:** glicogênio muscular; hipoxantinas; insulina; metabolismo; exercício.

<sup>1</sup> Instituto Superior e Centro Educacional Luterano Bom Jesus, Laboratório de Avaliação Multidisciplinar. R. Mafrá, 84, Bairro Saguacú, 89201-207, Joinville, SC, Brasil. Correspondência para/Correspondence to: A.E. LIMA-SILVA. E-mail: <limasilvae@hotmail.com>.

<sup>2</sup> Universidade do Estado de Santa Catarina, Centro de Educação Física e Desportos, Laboratório de Pesquisa Morfo-Funcional. Florianópolis, SC, Brasil

<sup>3</sup> Universidade Estadual de Londrina, Centro de Educação Física e Desportos, Grupo de Estudo das Adaptações Fisiológicas ao Treinamento. Londrina, PR, Brasil.

## ABSTRACT

*A large number of studies have been conducted to understand muscle glycogen metabolism during exercise. Classical studies demonstrated a relationship between the pre-exercise muscle glycogen content and duration of exercise. Muscle glycogen declines in a semilogarithmic manner in function of time, but glycogen concentration does not reach zero, which suggests that other fatigue mechanisms participate in the interruption of prolonged exercise. In this type of activity, glycogen depletion occurs first in slow twitch fibers followed by fast twitch fibers. The decrease in the rate of muscle glycogen utilization is synchronized with an increased rate of fat uptake, but the physiological mechanism is not well understood. Recent studies suggest that the decline of insulin during exercise could be a limiting factor of glucose transport through the plasma membrane, which increases the uptake of fatty acids. Others studies have also demonstrated that the structure of muscle glycogen itself can regulate the cellular uptake of free fatty acids via protein kinase. Physically, the glycogen molecule has two forms, one with a smaller molecular structure (approximately  $4 \cdot 10^5$  Da, proglycogen) and another one with a larger molecular structure (approximately  $10^7$  Da, macroglycogen). Apparently, the proglycogen form is more metabolically active during exercise and the macroglycogen form is more susceptible to increase with supercompensation diets. Higher concentrations of hypoxanthines and ammonia during exercise with muscle glycogen depletion have been reported, but studies that control exercise intensity better are necessary to help shed light on this issue.*

**Indexing terms:** muscle glycogen; hypoxanthines; insulin; metabolism; exercise.

## INTRODUÇÃO

O metabolismo energético durante o exercício, em especial do glicogênio muscular, tem sido amplamente investigado<sup>1-3</sup>. Bergstrom et al.<sup>1</sup> demonstraram que o tempo de sustentação de determinado exercício está relacionado com a quantidade de glicogênio muscular disponível para ressíntese da molécula de adenosina trifosfato (ATP). Nesse estudo verificou-se que níveis aumentados de glicogênio muscular, obtidos por combinação exercício-dieta (supercompensação), prorrogam o tempo de permanência no esforço, enquanto níveis reduzidos por jejum ou reposição inadequada de carboidratos dietéticos levam a uma diminuição no tempo de atividade. A partir desses achados, técnicos, treinadores e nutricionistas passaram a utilizar estratégias dietéticas para aumentar as reservas desse substrato.

Com o prolongamento do exercício, as reservas de glicogênio muscular diminuem progressivamente e parte da energia despendida no esforço passa a ser fornecida pelos triglicerídeos musculares, por glicose e por ácidos graxos livres (AGL) circulantes no plasma<sup>4</sup>. Entretanto, o conhecimento acerca dos mecanismos bioquímicos e fisiológicos que controlam a alternância dos substratos energéticos predominantes é limitado.

Estudos recentes sugerem que uma combinação entre ação hormonal (adrenalina, noradrenalina e insulina) e a própria estrutura molecular do glicogênio muscular regulam a entrada de substratos na fibra muscular<sup>5,6</sup>.

Diante do exposto, a intenção deste trabalho foi levantar as principais teorias envolvidas no metabolismo de glicogênio muscular durante o exercício. Serão discutidas a ação hormonal na regulação metabólica e a estrutura química do glicogênio muscular. Os estudos que demonstram associação entre o metabolismo de glicogênio muscular e formação de compostos bioquímicos (hipoxantinas e amônia) também serão debatidos. Quando necessário, será abordada a interação entre o metabolismo de carboidrato e de gordura.

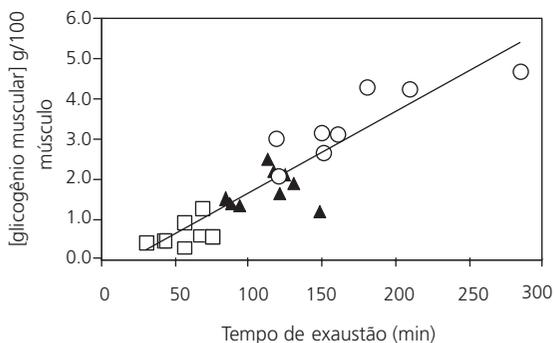
### Metabolismo do glicogênio muscular

A musculatura esquelética e o fígado constituem os principais órgãos de armazenamento de glicogênio. Embora encontremos no fígado uma maior concentração desse composto (até 6%), as reservas são maiores, em termos absolutos, na musculatura esquelética.

O metabolismo do glicogênio muscular durante o exercício foi elucidado a partir de uma

série de estudos clássicos, publicados pelo grupo do Instituto Karolinska de Estocolmo<sup>1,7,8</sup>. Esses estudos constituíram a base atual do conhecimento sobre o metabolismo do glicogênio muscular durante o exercício<sup>9</sup>, sendo utilizados na sustentação da maior parte das publicações subsequentes<sup>10-12</sup>. Entre os principais achados deste grupo estão: a correlação linear entre o tempo de fadiga em uma determinada intensidade ( $\%VO_{2max}$ ) e as concentrações iniciais de glicogênio no músculo (Figura 1), bem como a redução dos estoques de glicogênio (g/100g músculo seco) de forma semi-logarítmica em função do tempo, tendendo a se aproximar de zero no mesmo instante em que passa a ser difícil a manutenção da intensidade do exercício.

Estudos posteriores confirmaram esse comportamento de redução do glicogênio muscular em função do tempo de exercício<sup>2,10</sup>. Entretanto, destacou-se que a curva de glicogênio *versus* o tempo de exercício poderia apresentar comportamento trifásico, ou seja, um rápido declínio inicial, seguido por uma queda constante, e, finalmente, uma degradação mais lenta nos minutos finais. Inicialmente, a explicação para esse comportamento baseou-se na existência de um estado de hipóxia relativa nos momentos iniciais do exercício, levando a uma rápida degradação do glicogênio muscular, com conseqüente formação



**Figura 1.** Relação entre a concentração inicial de glicogênio muscular e o tempo de *performance*. □ após dieta baixa em carboidrato, ▲ após dieta balanceada e ○ após dieta alta em carboidrato.

Fonte: adaptado de Bergstrom et al.<sup>1</sup>.

de lactato sanguíneo. Na parte intermediária da curva, a queda constante poderia ser derivada da estabilização nos processos metabólicos, com equilíbrio entre a utilização do glicogênio muscular de forma aeróbia e a produção de lactato. Na última parte da curva, os estoques reduzidos de glicogênio muscular levariam a uma lenta degradação, aumentando gradativamente a utilização de gordura e glicose sangüínea como fonte de energia<sup>8</sup>.

Contudo, a primeira explicação atualmente pode ser contestada, pois um estado de hipóxia nem sempre é encontrado nas células musculares a ponto de impossibilitar a utilização do metabolismo oxidativo. A degradação rápida do glicogênio muscular, tendo como produto final o lactato, pode ser decorrente da ineficiência dos sistemas de transporte de íons  $H^+$  para dentro da mitocôndria, isto é, lançadeira glicerol-fosfato<sup>13</sup>, ou de uma inerente inércia das enzimas mitocondriais, responsáveis pelos processos oxidativos<sup>14</sup>.

A idéia de que as reservas de glicogênio muscular aproximam-se de zero no instante em que iniciam os sintomas de fadiga foram falseadas<sup>1,2,3,7,10,15</sup>. Em todos os estudos subsequentes a 1967, observou-se resquício de glicogênio muscular (~24%) ao final de exercícios prolongados (~70%  $VO_{2max}$ ), interrompidos pela exaustão do indivíduo<sup>2,15,16</sup>. Embora, na maior parte desses estudos, a fadiga tenha sido associada com a redução do glicogênio muscular, a pequena reserva restante seria suficiente para o prolongamento da atividade, sugerindo o envolvimento de outros mecanismos na interrupção de exercícios com essas características.

Outra série de estudos, iniciada a partir da década de 70, em Estocolmo, pelo grupo liderado por Gollnick, complementou o conhecimento sobre depleção de glicogênio muscular. Utilizando a técnica de histoquímica qualitativa, denominada "periódica reação do ácido-Schiff" (*periodic acid-Schiff's reaction - PAS*), Gollnick et al.<sup>17</sup> verificaram que, após a redução do glicogênio muscular com manipulação exercício-dieta, seguida por três dias

de dieta rica em carboidratos (~2.000kcal de carboidratos), as reservas de glicogênio muscular aumentavam, aproximadamente, 60% em relação a uma dieta mista. Ao final de 30 minutos de exercício na bicicleta ergométrica (74%  $VO_{2max}$ ), a concentração de glicogênio muscular foi maior quando uma dieta rica em carboidratos precedia o teste. A novidade desse trabalho foi a apresentação dos resultados de depleção seletiva do glicogênio muscular, ou seja, a maior parte do glicogênio utilizado advinha das fibras de contração lenta. Resultados similares também foram encontrados após uma corrida de 30km, embora a influência da dieta anterior ao exercício não tenha sido estudada<sup>18</sup>.

Em estudo posterior, Gollnick et al.<sup>19</sup> observaram que o glicogênio das fibras de contração rápida era o primeiro a ser depletado, após a realização de 6 séries de um minuto de duração (150% da potência aeróbia máxima), intercaladas por períodos de 10 minutos de repouso entre as séries. Em outra publicação, Gollnick et al.<sup>20</sup> finalizaram o modelo de depleção seletiva do glicogênio muscular, analisando diferentes intensidades de exercício na bicicleta ergométrica entre 30% a 150% do  $VO_{2max}$ . Os autores descobriram que a depleção do glicogênio muscular era 7,4 vezes maior a 84% do  $VO_{2max}$  do que a 31% do  $VO_{2max}$ , e a depleção era mais significativa nas fibras de contração lenta. Porém, com o prolongamento da atividade, um progressivo decréscimo no glicogênio muscular era observado também em fibras de contração rápida. Nos exercícios de intensidade acima do  $VO_{2max}$ , o glicogênio de ambas as fibras era depletado. Esses achados foram confirmados em estudos posteriores<sup>21</sup>.

Apesar do conhecimento obtido nas décadas de 60 e 70, os mecanismos fisiológicos e bioquímicos envolvidos na regulação da degradação do glicogênio muscular durante o exercício não foram totalmente esclarecidos. Em exercícios submáximos (entre 65-75%  $VO_{2max}$ ), a degradação (absoluta) do glicogênio diminui com o prolongamento da atividade, enquanto os AGL circulantes no plasma e a glicose sanguínea

aumentam sua participação na ressíntese do ATP. Isso parece mais evidente quando os níveis de glicogênio muscular pré-exercício encontram-se abaixo do normal<sup>22</sup>. Entretanto, pouco se sabe sobre os mecanismos que controlam essas alterações.

Uma questão inerente é: qual desses dois substratos tem preferência na “substituição” do glicogênio muscular? Um elegante trabalho de Weltan et al.<sup>5</sup> permite levantar algumas especulações (Tabela 1). Nesse estudo, os indivíduos foram designados aleatoriamente para um de quatro grupos, sendo que em três grupos a glicemia sanguínea foi mantida estável (euglicemia) através de infusão intravenosa de glicose. Desses três, um apresentava concentração inicial de glicogênio muscular normal, um depleção de glicogênio prévia e um depleção de glicogênio prévia mais infusão de insulina no exercício. No quarto grupo, também com depleção prévia de glicogênio, a taxa de infusão de glicose foi aumentada, a fim de manter uma situação de hiperglicemia. Os resultados demonstraram que a utilização do glicogênio muscular foi significativamente reduzida, nos grupos com depleção prévia de glicogênio. Além disso, foi verificado que, mesmo mantendo a glicemia sanguínea estável, o ácido graxo foi o substrato energético preferencialmente utilizado na situação de depleção de glicogênio, exceto quando foi mantida uma hiperglicemia ou hiperinsulinemia. Nessas duas últimas situações, a glicose sanguínea foi utilizada predominantemente como fonte energética.

**Tabela 1.** Alteração na taxa de utilização do glicogênio muscular, gordura e glicose, durante 145 minutos de cicloergômetro a 70% do consumo máximo de oxigênio, em três diferentes situações comparadas à situação de glicogênio muscular normal.

Glicogênio muscular	Taxa de utilização		
	GM	Gordura	Glicose
Euglicemia	Redução	Aumento	Semelhante
Euglicemia + insulina	Redução	Semelhante	Aumento
Hiperglicemia	Redução	Semelhante	Aumento

GM: glicogênio muscular; DG: depleção prévia de glicogênio muscular. Fonte: adaptado de Weltman et al.<sup>5</sup>.

Esses achados, embora se distanciem de situações fisiológicas normais, sugerem que em situações de depleção de glicogênio muscular, o músculo ativo utiliza, preferencialmente, os lipídios como substrato energético. Esse processo ocorre, provavelmente, por controle da noradrenalina, pois sua concentração no sangue aumenta significativamente com o exercício, elevando, assim, as concentrações de AGL plasmáticos e auxiliando na manutenção da glicemia sangüínea. Em contrapartida, a insulina em excesso (por infusão ou hiperglicemia) exerce efeitos antagônicos, estimulando o consumo de glicose pelo músculo, e inibindo a lipólise. A preferência para a utilização de lipídios como fonte de energia, na ausência de concentração adequada de glicogênio muscular, tem sido sustentada na literatura<sup>4</sup>.

O mecanismo fisiológico de restrição no consumo de glicose plasmática em situações de depleção de glicogênio muscular não está totalmente esclarecido, mas a hipótese mais provável seria a limitação no transporte de glicose através da membrana da célula. Estudos de Hespel & Richter<sup>23</sup> com animais demonstram que, ratos com depleção de glicogênio aumentam o transporte de glicose pela membrana em 25% durante 15 minutos de contração isométrica máxima, quando comparados a ratos com supercompensação (combinação de exercício e dieta). Porém, esse aumento não foi suficiente para restabelecer o metabolismo de carboidratos, sendo necessário um aumento concomitante no consumo de lipídios e aminoácidos, conforme já relatado em humanos<sup>24,25</sup>. Como discutido anteriormente, o aumento exógeno de insulina no exercício pode facilitar o transporte de glicose através da membrana plasmática e restabelecer o metabolismo de carboidrato nas situações de depleção de glicogênio muscular<sup>5</sup>.

Se, por um lado, a escolha da célula muscular em utilizar lipídios como fonte de energia proporciona uma "economia" de carboidratos, em especial de glicose sangüínea, por outro, constitui uma manobra que, inevitavelmente, prejudica a manutenção da intensidade do exercício e o

desempenho<sup>26</sup>. Isso porque os ácidos graxos necessitam de maior quantidade de oxigênio para serem oxidados. A quantidade de energia liberada por litro de oxigênio e a velocidade de degradação da molécula é maior quando a glicose é metabolizada, ao invés de ácidos graxos, justificando porque em exercício de intensidade elevada ( $\sim 85\% \text{VO}_{2\text{max}}$ ) os carboidratos são, preferencialmente, utilizados<sup>4,27</sup>. Como é importante para atletas, em competições de longa duração, realizar a prova na maior intensidade relativa possível, os carboidratos acabam constituindo a principal fonte de energia. Esse pensamento está de acordo com o conceito de *crossover* do metabolismo, que estabelece uma modificação da predominância de lipídios para carboidratos, acontecendo próximo a  $80\% \text{VO}_{2\text{max}}$ , e com pouca interferência do nível de aptidão aeróbia<sup>27</sup>.

A partir dos dados experimentais levantados nesta sessão, demonstrando que o consumo de glicose pelo músculo ativo não é significativamente aumentado em condições de depleção de glicogênio muscular, é possível imaginar que, de forma defensiva e prioritária, o organismo privilegie a oferta de glicose ao sistema nervoso central (SNC), protegendo-o de possíveis "lesões" por deficiência de nutriente. Esse raciocínio está de acordo com a hipótese de um "governador central" controlando os mecanismos de fadiga<sup>28</sup>.

### **Efeito da insulina e do exercício no transporte celular de glicose**

Conforme o discutido anteriormente, a elevação da glicemia sangüínea em conjunto com o excesso de insulina exógena, aumenta o catabolismo da glicose<sup>5</sup>. Tem sido documentado que, durante o exercício, a glicose é utilizada pela célula da fibra muscular de forma independente da insulina, provavelmente, por aumento no número de transportadores de membrana ativos, isto é, GLUT-4<sup>29</sup>. O mecanismo envolvido não está totalmente esclarecido, mas uma possível via seria a produção da 5'-AMP-ativador da proteína quinase (PKA), o qual aumentaria a expressão

gênica do GLUT-4<sup>30</sup>. Entretanto, apesar dessa ativação independente durante o exercício, o transporte pode permanecer parcialmente controlado pela insulina. Dados recentes de Christ-Roberts et al.<sup>31</sup> suportam essa concepção, demonstrando que o exercício com duração de 30 minutos a 70%  $VO_{2max}$  com infusão de insulina, aumenta o transporte de glicose pela membrana plasmática. Esse aumento no transporte foi devido a uma maior ligação entre o substrato 1 do receptor da insulina (IRS-1) com a subunidade PI3-quinase, desencadeando um potente efeito cascata no citoplasma, com subsequente fosforilação dos receptores serina/treonina quinase (PKB) e aumento na translocação do GLUT-4.

Aparentemente, a combinação insulina-exercício exerce um efeito amplificador, com maior consumo de glicose pela célula muscular. Provavelmente, apenas combinando exercício e infusão de insulina com manutenção exógena da glicose é que a predominância energética em exercícios prolongados passaria de lipídios para carboidratos.

Estudo de Nielsen et al.<sup>30</sup> demonstrou que, em exercício a 80%  $VO_{2max}$ , a fosforilação da subunidade catalítica  $\alpha$ -PKA (Thr<sup>172</sup>) foi menor em um grupo de pessoas treinadas aerobiamente, quando comparado com um grupo de sedentários. Interessante que a concentração de glicose plasmática foi aumentada (5,9, desvio-padrão - DP= 0,5 vs 4,7, DP= 0,3 mM) no grupo treinado. Esses resultados estão de acordo com os dados reportados por Coggan et al.<sup>32</sup>, que observaram, também a 80%  $VO_{2max}$ , uma menor taxa de desaparecimento (Rd) da glicose do sangue em ciclistas bem treinados, e com a sugestão de Richter et al.<sup>33</sup> de que a translocação do GLUT4 é menor em indivíduos treinados para a mesma carga absoluta. Nielsen et al.<sup>30</sup> mostraram, igualmente, uma utilização de glicogênio muscular similar entre os grupos, mesmo com os treinados apresentando um maior estoque inicial. Isso está de acordo, pelo menos em parte, com o conceito de *crossover*<sup>28</sup> e sugere um mecanismo poupador de glicose plasmática em indivíduos treinados aerobiamente.

Frente aos dados apresentados, três importantes constatações podem ser destacadas, referentes ao metabolismo de glicogênio muscular: 1) há utilização preferencial de AGL em situações de depleção de glicogênio; 2) o consumo de glicose pelo músculo esquelético durante o exercício pode ser transporte-limitado e; 3) quando os estoques de glicogênio estão em níveis normais, principalmente nas intensidades mais elevadas, existe uma preferência por essa fonte de energia. Assim, a glicose sangüínea vem a ser, então, um rico e precioso combustível, que deve ser utilizado, predominantemente, pelo músculo ativo, quando alta concentração plasmática desse substrato possa ser mantida (por infusão ou ingestão).

### Estrutura funcional do glicogênio muscular

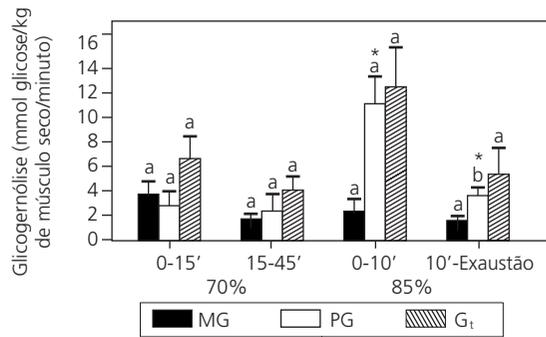
Mesmo com o integrado controle, apresentado na sessão anterior, envolvendo sinalizadores e receptores de membrana, a própria estrutura do glicogênio parece contribuir também para esse controle. Uma proposta de auto-regulação, a partir da integração física e enzimática da molécula de glicogênio, foi, recentemente, apresentada por Shearer & Graham<sup>6</sup>. O modelo foi elaborado a partir do desenvolvimento de um método semi-quantitativo de determinação do glicogênio muscular, utilizando a técnica de microscopia de transmissão eletrônica, a qual permite conhecer o número, a distribuição e a área de cada grânulo de glicogênio. Inicialmente, o grânulo cresce em um formato linear, com aumento seqüencial de unidades de glicose, sendo a primeira ligação unida à glicogenina, uma proteína auto-glicosilante. A partir dessa primeira ligação, mais unidades de glicose podem ser adicionadas pela ação de duas enzimas-chave no processo de síntese, a glicogênio sintetase (GS) e a enzima ramificadora. O acréscimo seqüencial e ramificado de glicose realizado por essas duas enzimas faz com que os estoques de carboidratos dentro da célula aumentem de forma exponencial. A molécula passa, então, a apresentar um formato

esférico, com o seu crescimento sendo inibido quando atinge um diâmetro de, aproximadamente, 42 nanômetros.

Existem duas principais formas de armazenamento do glicogênio, as quais podem ser identificadas pela sua solubilidade em ácido perclórico, denominadas proglicogênio (PG) e macroglicogênio (MG). A primeira constitui uma molécula menor ( $\sim 4 \cdot 10^5$  Da) com maior razão proteína/carboidrato. A segunda, uma molécula maior ( $\sim 10^7$  Da), com a mesma quantidade de proteína da PG, mas com mais carboidrato (menor razão proteína/carboidrato).

Um estudo de Graham et al.<sup>34</sup> demonstrou que PG e MG apresentam dinâmicas de degradação diferentes, que podem ser dependentes da intensidade de esforço. Nesse estudo, os autores demonstraram que a taxa de degradação de MG e PG era similar em exercícios a 70%  $VO_{2max}$ , mas muito maior de PG quando o exercício era realizado a 85%  $VO_{2max}$ . Em exercício intermitente (3 x 3 minutos a 100%  $VO_{2max}$ ), na primeira série as duas formas são utilizadas em proporções similares, mas com tendência a manter a preferência pela PG. Na segunda série, um declínio no metabolismo de MG e a manutenção de PG foram observados, apesar de não diferirem estatisticamente. Na última, as duas formas diminuíram a taxa de degradação. Os autores concluíram que a forma PG pode ser metabolicamente mais ativa e o metabolismo de MG pode ser rapidamente inibido com o passar do tempo, tanto em exercícios continuados, quanto repetidos. Esse fato não parece ser generalizável, uma vez que, como demonstrado na Figura 2 extraída do trabalho de Graham et al.<sup>34</sup>, é perceptível que a 85%  $VO_{2max}$  ocorre uma queda significativa na taxa de degradação do glicogênio, em função do tempo, apenas para PG, com MG mantendo-se constante. Resta ainda, em futuros estudos, determinar a mudança na degradação de MG e PG em exercícios prolongados, em especial de exercícios realizados até a exaustão.

Um outro estudo do mesmo grupo também mostrou que, após exercício a 70%  $VO_{2max}$



**Figura 2.** Taxa de glicogenólise (unidades de glicose em mmol/kg de músculo seco/minuto) do macroglicogênio (MG), proglicogênio (PG) e glicogênio total (G<sub>t</sub>) no início (0-15 ou 0-10 minutos) e no final (15-45 minutos ou 10 minutos-exaustão) de exercício, a 70% e 85% do  $VO_{2max}$ .

Fonte: Graham et al.<sup>34</sup>, utilizado com permissão.

realizado até a exaustão, seguido por uma dieta composta por 75% de carboidrato durante dois dias, a forma MG aumenta somente nas 24 e 48 horas após o exercício. Outro dado interessante é que, apesar de aumentar a forma macro, a micro (PG) é mantida em valores próximos a 350mmol unidades de glicose/kg de peso seco, muito semelhante à concentração total de glicogênio normal, sem supercompensação<sup>11</sup>. Aparentemente, o glicogênio "extra" foi armazenado na forma MG, mas preservando a concentração fisiológica de PG. Adamo et al.<sup>11</sup> sugerem que esses resultados, junto com os obtidos no estudo de Graham et al.<sup>34</sup>, são fortes indicadores de que o PG é metabolicamente mais ativo. Entretanto, algumas lacunas deixadas em aberto merecem ser mais bem investigadas, como a questão de se o aumento no tempo de exaustão causado pela supercompensação está associado à forma MG de glicogênio, e a averiguação dos possíveis efeitos da depleção prévia de glicogênio muscular sobre o metabolismo de MG e PG.

Outro ponto interessante é que o próprio metabolismo de glicogênio muscular pode se auto-regular. Conforme a "árvore" de glicogênio vai perdendo o conteúdo de glicose das extremidades, uma diminuição na atividade da enzima catalítica glicogênio fosforilase (GF) e o aumento na

atividade da enzima GS são observados<sup>6</sup>. O mecanismo exato da interferência física da molécula de glicogênio na ação enzimática não está muito bem explicado, mas assume-se que está relacionado a uma maior ativação da PKA em situações de diminuição na reserva de glicogênio, uma vez que essa enzima apresenta um sítio de ligação com a molécula de glicogênio e outro com a GS e GF. A PKA é uma importante proteína responsável pelo transporte de glicose e AGL para dentro da célula, o que poderia aumentar a oxidação e diminuir a síntese de substratos. Assim, a entrada de substrato na célula pode ser mediada parcialmente pelo conteúdo de glicogênio muscular.

Para a glicose 1-fosfato, liberada do glicogênio muscular, iniciar na via glicolítica, é necessário a conversão para glicose-6-fosfato, pela enzima fosfoglicomutase. Parece razoável imaginar que uma diminuição da atividade da GF acarretaria em menor formação de glicose 1-fosfato, e, conseqüentemente, de glicose 6-fosfato. A diminuição na concentração de glicose 6-fosfato no músculo esquelético é um potente estimulador alostérico da hexoquinase, a enzima responsável pela fosforilação da glicose vinda do sangue. Como a reação glicose 6-fosfato  $\leftrightarrow$  frutose 6-fosfato apresenta-se em equilíbrio, a diminuição da glicose 6-fosfato leva a uma concomitante diminuição da frutose 6-fosfato. Como essa última enquadra-se como um regulador alostérico da fosfofrutoquinase (PFK), essa enzima pode diminuir sua atividade nas situações em que a concentração de frutose-6-fosfato decai. Assim, a predominância de glicose sanguínea entrando na via glicolítica pode ocasionar uma simultânea redução na velocidade de degradação, causada pela menor atividade da enzima PFK. Como conseqüência, a redução da glicose a piruvato aconteceria de forma mais lenta, facilitando a entrada dessa molécula dentro da mitocôndria, o que evitaria a formação de lactato. Isso pode explicar porque alguns estudos apresentam uma menor concentração de lactato em exercício

submáximo e máximo em situações de depleção de glicogênio muscular<sup>25</sup>. Realmente, existem evidências de que a formação de lactato a partir do glicogênio muscular é 10 vezes maior do que a partir da glicose<sup>35</sup>.

### **Efeito da intensidade do esforço no metabolismo de glicogênio muscular e alterações bioquímicas intracelulares**

Para uma determinada porcentagem do  $VO_{2max}$ , na maior parte dos estudos sobre depleção de glicogênio muscular foi observado que, nessa situação, ocorre uma diminuição significativa no tempo de manutenção do esforço<sup>1,15</sup>. Além de todas as alterações metabólicas levantadas até o momento contribuírem para a diminuição na *performance*, uma interessante hipótese seria a existência de um declínio nos intermediários do ciclo de Krebs, levando, conseqüentemente, a uma menor ressíntese de ATP pela via aeróbia. Seguindo essa linha de raciocínio, isso aumentaria a concentração de ADP intramuscular, estimulando a reação da mioquinase e causando um acúmulo de inosina monofosfato (IMP), com formação de amônia ( $NH_3$ ). Entretanto, essa afirmação foi refutada recentemente em trabalho de Baldwin et al.<sup>3</sup>, que não conseguiram observar decréscimo na soma de quatro intermediários do ciclo (citrato, isocitrato, fumarato e malato) - os quais representam 70% do total - após ~100 minutos de exercício com depleção prévia de glicogênio ou ~150 minutos com supercompensação prévia de glicogênio, em uma intensidade referente a 70%  $VO_{2max}$ . Nesse mesmo estudo, a soma do total de adenina nucleotídeo (chamada de TAN, que é igual a ATP + ADP (adenosina difosfato) + AMP (adenosina monofosfato) não foi alterada em nenhuma das situações.

Esse mecanismo continua intrigante porque existem resultados conflitantes na literatura. Por exemplo, Spencer & Katz<sup>36</sup> observaram que, após um exercício de duração de ~5,5 minutos na inten-

sidade correspondente a ~95%  $VO_{2max}$ , o acúmulo de IMP foi maior em situações de depleção prévia de glicogênio, quando comparado com supercompensação prévia. Apesar do acúmulo de frutose 6-fosfato ter sido menor com depleção prévia (o que levaria a uma inibição da PFK), a glicólise não foi alterada devido à compensação exercida pelo acúmulo de ADP e AMP livre na célula, que funcionam como ativadores alostéricos da PFK. Resultados opostos foram encontrados por Febbraio & Dancey<sup>12</sup>, em que um exercício realizado a ~65%  $VO_{2max}$  (93% do limiar de lactato) até a exaustão não causou um significativo aumento em IMP ou hipoxantinas, e também não reduziu o TAN. Apesar de uma significativa relação entre tempo de exaustão e uso de glicogênio muscular ter sido encontrada ( $r=0,95$ ;  $p<0,05$ ), a associação entre IMP e glicogênio muscular, no final do exercício, não foi significativa ( $r=0,73$ ;  $p>0,05$ ).

As diferenças entre os estudos podem ser devidas, simplesmente, à forma de controle da intensidade do exercício. Por exemplo, Broberg & Sahlin<sup>37</sup> encontraram resultados diferentes de Febbraio & Dancey<sup>12</sup>, associando o acúmulo progressivo de  $NH_3$  com o baixo nível de glicogênio muscular ocasionado pelo exercício. A intensidade utilizada, entretanto, foi muito semelhante (~67%  $VO_{2max}$ ), mas a forma de determiná-la foi muito diferente. No estudo de Febbraio & Dancey<sup>12</sup>, a intensidade foi estabelecida a partir do limiar de lactato, o que, de certa forma, individualiza a intensidade de esforço, uma vez que, uma porcentagem fixa, estabelecida unicamente a partir do  $VO_{2max}$ , como a utilizada no estudo de Broberg & Sahlin<sup>37</sup>, pode representar uma “carga interna” muito diferente entre os indivíduos<sup>38</sup>. Essas diferenças metodológicas podem submeter os indivíduos a diversos domínios fisiológicos, sendo que os mecanismos de fadiga podem ser totalmente diferenciados, quando comparadas intensidades referentes aos limiares de lactato. Isso explica, também, porque, dependendo da intensidade

estudada, a depleção de glicogênio pode ou não estar associada à fadiga<sup>15</sup>.

Evidências mostram que o exercício realizado acima do  $VO_{2max}$  (supra-máximo) parece ter uma dependência menor da disponibilidade inicial de glicogênio muscular. Em estudo de Vandenbergue et al.<sup>39</sup>, a 125%  $VO_{2max}$ , a supercompensação de glicogênio levou a um aumento de 56% na concentração muscular inicial desse composto, sem, no entanto, aumentar a tolerância ao esforço (~175 s), ou modificar o acúmulo de lactato e de pH sanguíneos. Resultados similares foram encontrados por Hargreaves et al.<sup>40</sup>, que não identificaram nenhum efeito da supercompensação de glicogênio muscular sobre a potência de pico, potência média e máximo déficit acumulado de oxigênio em exercício de 75 segundos (75 *all-out*). Entretanto, em atividades com exigência mista ou participação efetiva da capacidade láctica (aeróbio-anaeróbio com duração entre 3 a 10 minutos, isto é, próximo ao  $VO_{2max}$ ), a depleção de glicogênio muscular pode interferir significativamente no desempenho. Newsholme et al.<sup>35</sup> estimaram a quantidade de glicogênio muscular utilizada pela via aeróbia e anaeróbia, em uma corrida de 5 mil metros (~13min) e demonstraram que ambas podem consumir quase todo o glicogênio armazenado no músculo. Assumindo que essa estimativa esteja correta, a fadiga por depleção de glicogênio poderia acontecer antes do acúmulo excessivo de prótons no músculo. Estudos com o objetivo de determinar a intensidade a partir da qual as reservas de glicogênio muscular deixam de ser importantes para o desempenho devem ser conduzidos, principalmente, comparando esforços abaixo e acima do  $VO_{2max}$ .

Digno de nota, nem todo glicogênio intracelular exerce função de regenerar ATP para contração muscular. Uma importante e significativa parcela destina-se a manter o funcionamento da bomba de cálcio e interfere, apenas indiretamente, no processo de contração - relaxamento<sup>41,42</sup>. Alguns autores sugerem que, mesmo com glicogênio total intracelular suficiente para manter a

atividade muscular, a depleção dos depósitos próximos à bomba de cálcio pode ocorrer precocemente, impossibilitando a continuidade do exercício<sup>12</sup>. Apesar de evidências indiretas sugerirem a existência desse mecanismo em humanos<sup>43</sup>, infelizmente, dentro do nosso conhecimento, não existem estudos que possam comprovar essa hipótese.

### **Reposição do glicogênio muscular e estratégias de supercompensação**

Embora não tenha sido o escopo principal dessa revisão, a compreensão das estratégias nutricionais de ressíntese do glicogênio muscular é de suma importância para o processo de recuperação de atletas em fase competitiva e pré-competitiva. A porcentagem de carboidratos em uma dieta balanceada comum aproxima-se de 60% do valor energético total, mas para aumentar as reservas de glicogênio muscular pré-competição, a porcentagem de carboidratos nos três dias que precedem a competição deveria aproximar-se de 80%<sup>39</sup>. O índice glicêmico do alimento<sup>44</sup> e o tipo de monossacarídeo utilizado<sup>45</sup> são importantes variáveis que precisam ser levadas em consideração. Stevenson et al.<sup>44</sup> observaram que o aumento da glicose sanguínea aos 30 e 120 minutos após o término do exercício era acentuado quando alimentos de alto índice glicêmico eram utilizados na reposição de carboidratos. O pico de insulina após 120 minutos do fim do exercício também foi maior após a ingestão de alimentos de alto índice glicêmico. Os autores sugeriram que o maior nível de insulina poderia aumentar a síntese de glicogênio muscular. Alguns autores demonstraram que a inclusão de proteínas na refeição pós-esforço acelera a reposição do glicogênio muscular<sup>46</sup>, mas nenhum efeito dessa estratégia foi observado por Wojcik et al.<sup>47</sup>, comparando suplementação de carboidrato com proteína + carboidrato. Portanto, a eficácia da ingestão de proteínas em conjunto com carboidratos, sobre a reposição do glicogênio muscular, é uma questão ainda não esclarecida.

Conlee et al.<sup>45</sup>, utilizando um modelo animal, constataram que o uso de frutose nas duas primeiras horas após o término do esforço ou jejum prolongado (24 horas) não aumenta significativamente as reservas de glicogênio muscular. Por outro lado, a ingestão de glicose aumenta consideravelmente as reservas de glicogênio muscular em ambas as situações (jejum ou exercício). Interessante que a frutose provocou um aumento na taxa de ressíntese do glicogênio hepático, quando comparada com a glicose. Além disso, a taxa de restauração do glicogênio hepático foi maior após o jejum do que após o exercício. Esses resultados sugerem que a frutose tenha uma maior importância no restabelecimento das reservas hepáticas de glicogênio, enquanto a glicose, na ressíntese do glicogênio muscular. Contudo, recomenda-se cautela ao aplicar esse modelo de reposição de glicogênio em humanos, uma vez que o mesmo foi testado apenas em animais.

Por fim, um interessante estudo de Lambert et al.<sup>48</sup> demonstrou que uma dieta rica em gordura (>65% MJ de gordura) durante 10 dias, seguida por 3 dias de dieta rica em carboidratos (>70% MJ de carboidrato), diminuiu significativamente a utilização de glicogênio muscular e o tempo necessário para percorrer 20km no ciclismo. Essa comparação foi feita em relação a um procedimento controle com a ingestão de uma dieta mista (~30% MJ de gordura) nos 10 dias anteriores aos três dias de sobrecarga de carboidrato. Esse estudo abre um novo campo de investigação referente a possíveis combinações de dieta, como estratégia para aumentar as reservas de glicogênio muscular pré-competição e melhorar o desempenho esportivo.

A partir desses achados, fica clara a importância da reposição de carboidratos após o exercício. Uma dieta rica em carboidratos (~80% do valor energético total) com alto teor de glicose após o exercício prolongado, deveria ser aplicada para a ressíntese mais efetiva do glicogênio muscular e recuperação do atleta.

## CONSIDERAÇÕES FINAIS

As reservas de glicogênio muscular são estreitamente relacionadas ao desempenho e tempo de sustentação do esforço em determinado exercício. A transferência de predominância do metabolismo de glicogênio muscular para o de lipídios acontece com o prolongamento da atividade, à medida que diminuem as reservas de carboidrato. O mecanismo fisiológico que limita a entrada de glicose na fibra muscular ainda não está totalmente esclarecido, mas, provavelmente, um balanço entre a diminuição da insulina com o exercício e a modificação física na estrutura molecular do glicogênio muscular regule esse processo. Essa limitação no transporte de glicose pode prevenir o estado de hipoglicemia, por poupar a utilização desse substrato pelo músculo. Algumas evidências apontam para uma relação entre glicogênio muscular e bomba de cálcio no processo de contração - relaxamento, mas estudos em humanos são necessários para comprovar essa hipótese. Estudos com hipoxantinas e amônia também não permitem muitas conclusões, e desenhos experimentais, com melhor controle da intensidade do exercício, podem elucidar essa questão.

## COLABORADORES

A.E. LIMA-SILVA concebeu a idéia do trabalho, e desenvolveu a metodologia, a revisão da literatura e a redação. T.C. FERNANDES participou nas discussões referentes ao desenvolvimento da idéia do trabalho, auxiliou na análise crítica da literatura e na redação. F.R. OLIVEIRA participou nas principais discussões referentes à idéia do trabalho e no desenvolvimento da metodologia. Contribuiu significativamente com o modelo teórico e com a orientação do trabalho. F.Y. NAKAMURA participou nas principais discussões referentes à idéia do trabalho, auxiliou na revisão de artigos e na redação final do trabalho. M.S. GEVAERD participou nas principais discussões referentes à idéia do trabalho, na redação final e na orientação do trabalho.

## REFERÊNCIAS

1. Bergstrom J, Hermansen L, Hultman E, Saltin B. Diet, muscle glycogen and physical performance. *Acta Physiol Scand.* 1967; 71(2):140-50.
2. Bosch AN, Dennis SC, Noakes TD. Influence of carbohydrate loading on fuel substrate turnover and oxidation during prolonged exercise. *J Appl Physiol.* 1993; 74(4):1921-7.
3. Baldwin J, Snow RJ, Gibala MJ, Garnham A, Howarth K, Febbraio MA. Glycogen availability does not affect the TCA cycle or TAN pools during prolonged, fatiguing exercise. *J Appl Physiol.* 2003; 94(6):2181-7.
4. Romijn JA, Coyle EF, Sidossis LS, Gastaldelli A, Horowitz JF, Endert E, et al. Regulation of endogenous fat and carbohydrate metabolism in relation to exercise intensity and duration. *Am J Physiol. (Endocrinol Metab)* 1993; 265 (3 Pt 1): E380-91.
5. Weltan SM, Bosch AN, Dennis SC, Noakes TD. Influence of muscle glycogen content on metabolic regulation. *Am J Physiol. (Endocrinol Metab)* 1998; 274(1 Pt 1):E72-82.
6. Shearer J, Graham TE. Novel aspects of skeletal muscle glycogen and its regulation during rest and exercise. *Exerc Sports Sci Rev.* 2004; 32(3):120-6.
7. Ahlborg B, Bergstrom J, Ekelund L-G, Hultman E. Muscle glycogen and muscle electrolytes during prolonged physical exercise. *Acta Physiol Scand.* 1967; 70:129-42.
8. Bergstrom J, Hultman E. A study of the glycogen metabolism during exercise in man. *Scand J Clin Lab Invest.* 1967; 19(3):218-28.
9. Conlee RK. Muscle glycogen and exercise endurance: a twenty-year perspective. *Exerc Sport Sci Rev.* 1987; 15:1-28.
10. Coyle EF, Coggan AR, Hemmert MK, Ivy JL. Muscle glycogen utilization during prolonged strenuous exercise when fed carbohydrate. *J Appl Physiol.* 1986; 61(1):165-72.
11. Adamo KB, Tarnopolsky MA, Graham TE. Dietary carbohydrate and postexercise synthesis of proglycogen and macroglycogen in human skeletal muscle. *Am J Physiol (Endocrinol Metab).* 1998; 275(2 Pt 1):E229-34.
12. Febbraio MA, Dancy J. Skeletal muscle energy metabolism during prolonged, fatigue exercise. *J Appl Physiol.* 1999; 87(6):2341-7.
13. Stainsby WN. Biochemical and physiological bases for lactate production. *Med Sci Sports Exerc.* 1986; 18(3):341-3.
14. Grassi B, Gladden LB, Samaja M, Stary CM, Hogan MC. Faster adjustment of O<sub>2</sub> delivery does not

- affect  $\text{VO}_2$  on-kinetics in isolated in situ canine muscle. *J Appl Physiol*. 1998; 85(4):1394-403.
15. Loy SF, Conlee RK, Winder WW, Nelson AG, Arnall DA, Fisher AG. Effects of 24-hour fast on cycling endurance time at two different intensities. *J Appl Physiol*. 1986; 61(2):654-9.
  16. Tsintzas OK, Williams C, Boobis L, Greenhaff P. Carbohydrate ingestion and single muscle fiber glycogen metabolism during prolonged running in men. *J Appl Physiol*. 1996; 81(2):801-9.
  17. Gollnick PD, Piehl K, Saubert ICW, Armstrong RB, Saltin B. Diet, exercise, and glycogen changes in human muscle fibers. *J Appl Physiol*. 1972; 33(4):421-5.
  18. Costill DL, Gollnick PD, Jansson ED, Saltin B, Stein EM. Glycogen depletion pattern in human muscle fibres during distance running. *Acta Physiol Scand*. 1973; 89(3):374-83.
  19. Gollnick PD, Armstrong RB, Sembrowich WL, Shepherd RE, Saltin B. Glycogen depletion pattern in human skeletal muscle fibers after heavy exercise. *J Appl Physiol*. 1973; 34(5):615-8.
  20. Gollnick PD, Piehl K, Saltin B. Selective glycogen depletion pattern in human muscle fibres after exercise of varying intensity and at varying pedalling rates. *J Physiol*. 1974; 241(1):45-57.
  21. Vollestad NK, Vaage O, Hermansen L. Muscle glycogen depletion patterns in type I and subgroups of type II fibres during prolonged severe exercise in man. *Acta Physiol Scand*. 1984; 122(4):433-41.
  22. Costill DL, Bowers R, Branam G, Sparks K. Muscle glycogen utilization during prolonged exercise on successive days. *J Appl Physiol*. 1971; 31(6):834-8.
  23. Hespel P, Richter EA. Glucose uptake and transport in contracting, perfused rat muscle with different pre-contraction glycogen concentrations. *J Physiol*. 1990; 427:347-59.
  24. Coggan AR. Plasma glucose metabolism during exercise in humans. *Sports Med*. 1991; 11(2):102-24.
  25. Blomstrand E, Saltin B. Effect of muscle glycogen on glucose, lactate and amino acid metabolism during exercise and recovery in human subjects. *J Physiol*. 1999; 514(Pt 1):293-302.
  26. Manetta J, Brun JF, Perez-Martin A, Callis A, Prefaut C, Mercier J. Fuel oxidation during exercise in middle-aged men: role of training and glucose disposal. *Med Sci Sports Exerc*. 2002; 34(3):423-9.
  27. Brooks GA, Mercier J. Balance of carbohydrate and lipid utilization during exercise: the "crossover" concept. *J Appl Physiol*. 1994; 76(6):2253-61.
  28. Noakes TD, Peltonen JE, Rusko HK. Evidence that a central governor regulates exercise performance during acute hypoxia and hyperoxia. *J Exp Biol*. 2001; 204(Pt 18):3225-34.
  29. Jessen N, Pold R, Buhl ES, Jensen LS, Schimtz O, Lund S. Effects of AICAR and exercise on insulin-stimulated glucose uptake, signaling, and GLUT-4 content in rat muscles. *J Appl Physiol*. 2003; 94(4):1373-9.
  30. Nielsen JN, Mustard KJW, Graham DA, Yu H, MacDonald CS, Pilegaard H, et al. 5'-AMP-activated protein kinase activity and subunit expression in exercise-trained human skeletal muscle. *J Appl Physiol*. 2003; 94(2):631-41.
  31. Christ-Roberts CY, Pratipanawat T, Pratipanawat W, Berria R, Belfort R, Mandarino LJ. Increased insulin receptor signaling and glycogen synthase activity contribute to the synergistic effect of exercise on insulin action. *J Appl Physiol*. 2003; 95(6):2519-29.
  32. Coggan AR, Raguso CA, Williams BD, Sidossis LS, Gastaldelli A. Glucose kinetics during high-intensity exercise in endurance-trained and untrained humans. *J Appl Physiol*. 1995; 78(3):1203-7.
  33. Richter EA, Wojtaszewski JF, Kristiansen S, Dugaard JR, Nielsen JN, Derave W, et al. Regulation of muscle glucose transport during exercise. *Int J Sport Nutr Exerc Metab*. 2001; 11(Suppl):S71-7.
  34. Graham TE, Adamo KB, Sheare J, Marchand I, Saltin B. Pro-and macroglycogenolysis: relationship with exercise intensity and duration. *J Appl Physiol*. 2001; 90(3):873-9.
  35. Newsholme EA, Blomstrand E, Ekblom B. Physical and mental fatigue: Metabolic mechanisms and importance of plasma amino acids. *Br Med Bull*. 1992; 48(3):477-95.
  36. Spencer MK, Katz A. Role of glycogen in control of glycolysis and IMP formation in human muscle during exercise. *Am J Physiol*. (Endocrinol Metab) 1991; 260(6 Pt 1): E859-64.
  37. Broberg S, Sahlin K. Hyperammonemia during prolonged exercise: an effect of glycogen depletion? *J Appl Physiol*. 1988; 65(6):2475-7.
  38. Weltman A, Weltman J, Rutt R, Seip R, Levine S, Snead D, et al. Percentages of maximal heart rate, heart rate reserve, and  $\text{VO}_2$  peak for determining endurance training intensity in sedentary women. *Int J Sports Med*. 1989; 10(3):212-6.
  39. Vanderberghe K, Hespel P, Eynde BV, Lysens R, Richter EA. No effect of glycogen level on glycogen metabolism during high intensity exercise. *Med Sci Sports Exerc*. 1995; 27(9):1278-83.

40. Hargreaves M, Finn JP, Withers RT, Halbert JA, Scroop GC, MacKay M, et al. Effect of muscle glycogen availability on maximal exercise performance. *Eur J Appl Physiol Occup Physiol*. 1997; 75(2):188-92.
41. Fridén J, Seger J, Ekblom B. Topographical localization of muscle glycogen: an ultrahistochemical study in human *vastus lateralis*. *Acta Physiol Scand*. 1989; 135(3):381-91.
42. Chin ER, Allen DG. Effects of reduced muscle glycogen concentration on force, Ca<sup>2+</sup> release and contractile protein function in intact mouse skeletal muscle. *J Physiol*. 1997; 498(Pt 1):17-29.
43. Booth J, McKenna MJ, Rueli PA, Gwinn TH, Davis GM, Thompson MW, et al. Impaired calcium pump function does not slow relaxation in human skeletal muscle after prolonged exercise. *J Appl Physiol*. 1997; 83(2):511-21.
44. Stevenson E, Williams C, Biscoe H. The metabolic responses to high carbohydrate meals with different glycemic indices consumed during recovery from prolonged strenuous exercise. *Int J Sport Nutr Exerc Metab*. 2005; 15(3):291-307.
45. Conlee RK, Lawler RM, Ross PE. Effects of glucose or fructose feeding on glycogen repletion in muscle and liver after exercise or fasting. *Ann Nutr Metab*. 1987; 31(2):126-32.
46. Berardi JM, Price TB, Noreen EE, Lemon PW. Postexercise muscle glycogen recovery enhanced with a carbohydrate-protein supplement. *Med Sci Sports Exerc*. 2006; 38(6):1106-13.
47. Wojcik JR, Walber-Rankin J, Smith LL, Gwazdauskas FC. Comparison of carbohydrate and milk-based beverages on muscle damage and glycogen following exercise. *Int J Sport Nutr Exerc Metab*. 2001; 11(4):406-19.
48. Lambert EV, Goedecke JH, Zyle C, Murphy K, Hawley JA, Dennis SC, et al. High-fat diet *versus* habitual diet prior to carbohydrate loading: effects of exercise metabolism and cycling performance. *Int J Sport Nutr Exerc Metab*. 2001; 11(2):209-25.

Recebido em: 12/5/2006

Versão final reapresentada em: 4/4/2007

Aprovado em: 17/5/2007