



**UNIVERSIDADE ESTADUAL DE  
CAMPINAS**  
**Instituto de Biologia**

**VALÉRIA SCORSATO**

**“Regulação de proteínas fosfatases envolvidas em câncer:  
estudos estruturais da TIPRL e interação com a  
Proteína Fosfatase 2A”**

**“Regulation of protein phosphatases involved in cancer:  
structural studies of TIPRL and its interaction with  
Protein Phosphatase 2A”**

**CAMPINAS**

**2015**

VALÉRIA SCORSATO

**“Regulação de proteínas fosfatases envolvidas em câncer:  
estudos estruturais da TIPRL e interação com a  
Proteína Fosfatase 2A”**

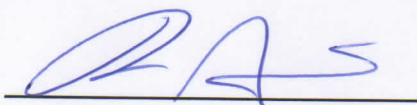
**“Regulation of protein phosphatases involved in cancer:  
structural studies of TIPRL and its interaction with  
Protein Phosphatase 2A”**

Tese apresentada ao Instituto de Biologia da  
Universidade Estadual de Campinas como parte dos  
requisitos exigidos para a obtenção do título de  
Doutora em Genética e Biologia Molecular, na área de  
Genética Animal e Evolução.

Thesis presented to the Institute of Biology the  
University of Campinas in partial fulfillment of the  
requirements for the degree of Doctor Genetics and  
Molecular Biology, in the area of Animal Genetics and  
Evolution

ORIENTADOR/SUPERVISOR: Prof. Dr. RICARDO APARICIO  
CO-ORIENTADOR/CO-SUPERVISOR: DRA. JULIANA HELENA COSTA SMETANA

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL  
DA TESE DEFENDIDA PELA ALUNA VALÉRIA SCORSATO  
E ORIENTADA PELO PROF. DR. RICARDO APARICIO.



CAMPINAS

2015

**Agência(s) de fomento e nº(s) de processo(s):** FAPESP, 2011/03054-9; CNPq, 141356/2011-1

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Biologia  
Mara Janaina de Oliveira - CRB 8/6972

Sco79r Scorsato, Valéria, 1985-  
Regulação de proteínas fosfatases envolvidas em câncer : estudos estruturais da TIPRL e interação com a proteína fosfatase 2A / Valéria Scorsato. – Campinas, SP : [s.n.], 2015.

Orientador: Ricardo Aparicio.  
Coorientador: Juliana Helena Costa Smetana.  
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Fosfatases. 2. Proteína fosfatase 2A. 3. Proteína TIPRL humana. 4. Interações proteína-proteína. 5. Câncer. 6. Cristalografia de proteínas. 7. Selenometionina. I. Aparicio, Ricardo, 1971-. II. Smetana, Juliana Helena Costa. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

#### Informações para Biblioteca Digital

**Título em outro idioma:** Regulation of protein phosphatases involved in cancer : structural studies of TIPRL and its interaction with protein phosphatase 2A

**Palavras-chave em inglês:**

Phosphatases  
Protein phosphatase 2A  
TIPRL protein, human  
Protein-protein interactions  
Cancer  
Protein crystallography  
Selenomethionine

**Área de concentração:** Genética Animal e Evolução

**Titulação:** Doutora em Genética e Biologia Molecular

**Banca examinadora:**

Ricardo Aparicio  
João Renato Carvalho Muniz  
Marcio Chain Bajgelman  
Celso Eduardo Benedetti  
Leandro Martínez

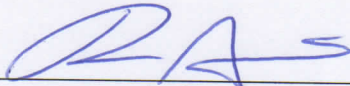
**Data de defesa:** 25-08-2015

**Programa de Pós-Graduação:** Genética e Biologia Molecular

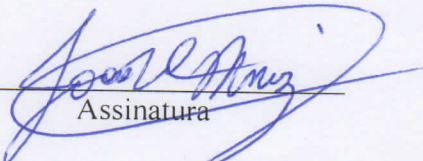
Campinas, 25 de agosto de 2015

**BANCA EXAMINADORA**

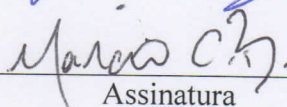
Prof Dr. Ricardo Aparicio (orientador)

  
Assinatura

Prof. Dr. João Renato Carvalho Muniz

  
Assinatura

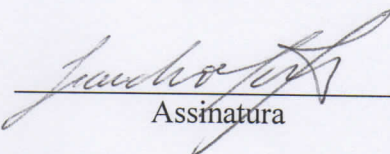
Dr. Marcio Chaim Bajgelman

  
Assinatura

Prof. Dr. Kleber Gomes Franchini

\_\_\_\_\_  
Assinatura

Dr. Leandro Martínez

  
Assinatura

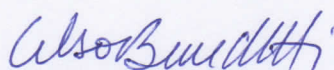
Dr. Adriano Rodrigues Azzoni

\_\_\_\_\_  
Assinatura

Dra. Melissa Regina Fessel

\_\_\_\_\_  
Assinatura

Dr. Celso Eduardo Benedetti

  
Assinatura

## DEDICATÓRIA

*Dedico esta Tese à minha família.*

## **AGRADECIMENTOS**

Ao meu orientador e amigo, Prof. Dr. Ricardo Aparicio, por ter acreditado no meu potencial como pesquisadora e aceitado, tão gentilmente, me orientar no desafio de um doutorado direto. Ficarei eternamente grata pelos ensinamentos, apoio e confiança na execução desta tese. Muito obrigada.

A minha coorientadora e amiga, Dra. Juliana H. C. Smetana, por me apresentar o incrível mundo da TIPRL, por ter aceitado me coorientar, pelos conselhos e por confiar no meu trabalho. Muito obrigada.

A minha família pela educação, pelas oportunidades e pelo amor que me deu suporte em cada nova etapa da minha vida.

Aos membros das bancas de qualificação e pré-banca – Adriano Azzoni, Goran Neshich, Jörg Kobarg, José Andrés Yunes e Melissa Fessel, pelas ótimas sugestões sobre o desenvolvimento deste trabalho.

Aos membros da banca desta tese, João Muniz, Kleber Franchini, Leandro Martínez e Marcio Bajgelman e os membros suplentes Adriano Azzoni, Celso Benedetti e Melissa Fessel pela disponibilidade e aceite do convite.

Aos colaboradores científicos Prof. Dr. Fábio Gozzo, Dr. Humberto Pereira, Dr. José Brandrão-Neto, Dr. Peter Schuck, Ma. Tatiani Lima e Ma. Germanna Righetto.

Ao Prof. Dr. Ronaldo Aloise Pilli pelo suporte e apoio e a todos os professores envolvidos no projeto temático: "Biologia Química: novos alvos moleculares naturais e sintéticos contra o câncer. Estudos estruturais, avaliação biológica e modo de ação" (processo FAPESP: 09/51602-5) e seus alunos que disponibilizaram compostos químicos para os ensaios de inibição com a proteína fosfatase PTP1B.

Aos colegas do grupo LABEc pela amizade, descontrações durante o período de desenvolvimento da tese e discussões: Dra. Emanuella Fonseca, Fábio Lazzarotto de Oliveira, Dra. Daniela Trivella, Dr. Marcelo Reis, Kishore Reddy Mandapati, Mariana Pereira Dias, Caroline Moreira Aguiar, Caio Nakavaki.

Aos colegas do LNBio pela amizade Ma. Ana Iacia, Ma. Nádia Rasheed, Vitor Amancio, Jhenifer Grunow e Amanda Eufrásio.

A Dra. Deborah de Alencar Simoni e Ma. Amanda Pocol pela amizade e apoio.

Aos funcionários e infraestrutura dos Institutos de Química e Biologia da UNICAMP e do Laboratório Nacional de Biociências.

Ao Laboratório Nacional de Luz Síncrotron (LNLS/CNPEM – Campinas SP) pelo apoio nas coletas de cristalografia de proteínas e SAXS e ao Diamond Light Source/Inglaterra.

Ao ROBOLAB e ao LEC (LNBio/CNPEM – Campinas SP) pelos ensaios de cristalização da proteína TIPRLAN e de DLS.

Ao colega Dr. José Sergio pela companhia em algumas coletas de dados de SAXS.

Ao Programa de Pós-Graduação em Genética e Biologia Molecular, ao coordenador Prof. Dr. Marcelo Menossi e à secretária Lourdes.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa no início deste projeto – processo nº 141356/2011-1.

A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela bolsa de doutorado direto – processo nº 2011/03054-9.

## RESUMO

O câncer é um conjunto de doenças de proliferação celular descontrolada cuja incidência na população mundial cresce continuamente ano após ano, atingindo índices alarmantes. A origem dos diversos tipos de câncer está em alterações genéticas e metabólicas das células, dentre as quais se destacam as alterações na sinalização celular mediada pela fosforilação reversível de proteínas, um processo altamente regulado por proteínas quinases e fosfatases. Neste trabalho, foram estudadas as fosfatases humanas PP2A e PTP1B. A proteína tirosina fosfatase PTP1B é superexpressa em tumores e está envolvida no câncer e em outras doenças como resistência a insulina e diabetes tipo 2, sendo considerada um bom alvo para o desenvolvimento de novas drogas. Nessa direção, foram analisadas várias séries de pequenas moléculas quanto à capacidade inibitória, tendo sido realizados esforços na obtenção de complexos cristalográficos. A PP2A é uma serina-treonina fosfatase que atua na célula como um importante supressor tumoral, exercendo sua atividade na forma de heterotrímeros formados por uma subunidade catalítica (PP2Ac), uma de ancoragem (PR65/A) e uma subunidade reguladora variável pertencente a uma dentre quatro famílias distintas (B/B'/B" e *striatins*). O processo de montagem desses heterotrímeros é fortemente regulado por outras proteínas, e envolve modificações pós-traducionais no extremo C-terminal da PP2Ac em um processo conhecido como biogênese da PP2A. A proteína TIPRL inibe a atividade de PP2A e há evidências de que atue em algum ponto do processo de biogênese. Esta tese reporta avanços significativos no estudo da TIPRL e sua interação com a PP2Ac, dentre eles, a estrutura cristalográfica da TIPRL, determinada a partir de proteína marcada com Se-Met, uma vez que não havia qualquer informação estrutural disponível nos bancos de dados, que pudesse ser utilizada para Substituição Molecular. A estrutura apresenta um novo tipo de enovelamento e um inesperado sítio de interação com o C-terminal da PP2A, conforme indicam nossos estudos, realizados por mutagênese sítio-dirigida, *pull-down* e análise de troca hidrogênio/deutério acoplada a espectrometria de massas. Esses resultados permitiram, pela primeira vez, propor uma função biológica para a TIPRL na biogênese da PP2A.



## ABSTRACT

The incidence of cancer in the world population continually grows year after year, reaching alarming levels. Cancer is characterized by uncontrolled cell proliferation, originated from genetic and metabolic anomalies including changes in cell signaling mediated by the reversible phosphorylation of proteins, a highly regulated process mediated by protein kinases and phosphatases. The present work focused on studies of human phosphatases PP2A and PTP1B. The protein tyrosine phosphatase PTP1B is overexpressed in tumors and is involved in cancer and other diseases such as insulin resistance and type 2 diabetes, and is considered a good target for the development of new drugs. Accordingly, we analyzed several series of small molecules for inhibitory capacity, and efforts have been made to obtain crystallographic structures of enzyme-drug complexes. PP2A is a serine-threonine phosphatase that functions as a major cellular tumor suppressor and exerts its activity as heterotrimers consisting of a catalytic subunit (PP2Ac), a scaffolding (PR65/A) and a variable regulatory subunit belonging to one of four distinct families (B/B'/B'' and striatins). The assembly process of such heterotrimers is tightly regulated by other proteins, and involves post-translational modifications at the C-terminus of PP2Ac in a process known as PP2A biogenesis. The TIPRL protein inhibits PP2A activity and there is evidence that this protein might function in the biogenesis process. This thesis reports significant advances in the study of TIPRL and its interaction with PP2Ac, such as the crystal structure of TIPRL determined from Se-Met labelled protein since there was no structural information available in databases that could be used for Molecular Replacement. The structure presents a new folding and an unexpected site binding site for the C-terminus of PP2A, which was validated by site-directed mutagenesis, pulldown and of hydrogen/deuterium exchange analysis by mass spectrometry. These results allowed us to propose a biological role for TIPRL in the biogenesis of PP2A.

# LISTA DE ILUSTRAÇÕES

<b>Figura 1.1:</b> Heterotrímero funcional da PP2A.....	25
<b>Figura 1.2:</b> Combinações para a formação do heterotrímero funcional da PP2A. ....	25
<b>Figura 1.3:</b> Esquema da biogênese do heterotrímero funcional da fosfatase PP2A.. ....	27
<b>Figura 1.4:</b> Estrutura da PTP-1B e elementos estruturais importantes para catálise e desenho de inibidores.....	32
<b>Figura 1.5:</b> Esquema da reação de remoção de grupo fosfato pela fosfatase PTP1B. ....	33
<b>Figura 2.1:</b> Esquema das construções das principais proteínas estudadas neste trabalho.....	39
<b>Figura 2.2:</b> Esquema do mecanismo de reação da PTP1B com o substrato sintético pNPP.....	46
<b>Figura 2.3:</b> Difratômetro Bruker Kappa APEX II Duo do Instituto de Química/Unicamp.....	49
<b>Figura 3.1:</b> Purificação da proteína tirosina fosfatase humana PTP1B.....	58
<b>Figura 3.2:</b> Ensaios da enzima PTP1B para adequação dos protocolos de inibição com compostos sintetizados pelos grupos de Química Orgânica que participaram do projeto temático.. ....	59
<b>Figura 3.3:</b> Gráfico dos inibidores da atividade catalítica da PTP1B.....	60
<b>Figura 3.4:</b> Cristais da proteína PTP1B e seu padrão de difração. ....	61
<b>Figura 3.5:</b> Mapa de densidade eletrônica da proteína PTP1B na presença do inibidor RPJ285 .....	65
<b>Figura 3.6:</b> Mapa de densidade eletrônica da proteína PTP1B na presença do inibidor PM02. ....	66
<b>Figura 3.7:</b> Mapa de densidade eletrônica da proteína PTP1B na presença do inibidor KSR1.. ....	66
<b>Figura 3.8:</b> Curvas de CD para os complexos de PP2A $\alpha$ 2. ....	68
<b>Figura 3.9:</b> Índícios de formação de cristais das amostras do complexo heterodímero. ....	69
<b>Figura 3.10:</b> Perfil de purificação da TIPRL $\Delta$ N por exclusão de peso molecular em coluna Superdex 200 16/60. ....	71
<b>Figura 3.11:</b> Cristais de TIPRL $\Delta$ N crescidos no experimento de cristalização do ROBOLAB e reproduzido no laboratório.....	72
<b>Figura 3.12:</b> Cristais de TIPRL $\Delta$ N crescidos em solução de SPG e seu padrão de difração.....	73
<b>Figura 3.13:</b> Modelos gerados computacionalmente para a Substituição Molecular. ....	74
<b>Figura 3.14:</b> Purificação e cristalização da TIPRL $\Delta$ N marcada com selenometionina. ....	75
<b>Figura 3.15:</b> Dimerização da TIPRL $\Delta$ N. ....	77
<b>Figura 3.16:</b> Cromatografia de exclusão de peso molecular para as amostras de TIPRL $\Delta$ N com e sem cauda de histidina e curvas de espalhamento de SAXS para a amostra TIPRL $\Delta$ N sem cauda. ....	79
<b>Figura 3.17:</b> Organização da TIPRL $\Delta$ N dentro do cristal. ....	86
<b>Figura 3.18:</b> Estrutura global da proteína TIPRL $\Delta$ N. ....	87
<b>Figura 3.19:</b> Cavidade da TIPRL $\Delta$ N onde abriga o peptídeo.....	88
<b>Figura 3.20:</b> Propriedades da superfície da TIPRL $\Delta$ N.....	89
<b>Figura 3.21:</b> Flexibilidade da estrutura da TIPRL $\Delta$ N.....	93
<b>Figura 3.22:</b> Mapa de cobertura da sequência de resíduos de aminoácidos da TIPRL <i>Full</i> obtidas pela otimização do ensaio de HDX.....	94
<b>Figura 3.23:</b> Representação de alguns gráficos de incorporação de deutério para a TIPRL <i>Full</i> na presença de peptídeos ligantes DYFL e DpYFL.....	95

<b>Figura 3.24:</b> Representação estrutural da análise HDX-MS da TIPRL <i>Full</i> com cauda de histidina na presença de peptídeos. ....	96
<b>Figura 3.25:</b> Mutagênese sítio-dirigida associada a GST- <i>pulldown</i> .. ....	99
<b>Figura 3.26:</b> Comparação estrutural entre o ácido chicórico e os peptídeos que interagem com a TIPRL. ....	100
<b>Figura 5.1:</b> Provável função da TIPRL na biogênese da PP2A.. ....	142

## LISTA DE TABELAS

<b>Tabela 1.1:</b> Lista de genes que codificam as isoformas das subunidades ABC para montagem do heterotrímero funcional da PP2A. ....	26
<b>Tabela 2.1:</b> Códigos de entrada do GeneBank e UniProt das sequências de nucleotídeos e aminoácidos das fosfatases PTP1B e PP2Ac e as proteínas TIPRL e $\alpha 4$ e os vetores onde os cDNA foram clonados. ....	39
<b>Tabela 2.2:</b> Sequências de aminoácidos das proteínas TIPRL $\Delta$ N e PTP1B.. ....	40
<b>Tabela 2.3:</b> Condições de expressão das proteínas recombinantes. ....	41
<b>Tabela 2.4:</b> Relação dos tampões utilizados para os processos de purificação de cada proteína. ....	43
<b>Tabela 2.5:</b> <i>Primers</i> utilizados para as mutações pontuais na TIPRL <i>Full</i> . ....	54
<b>Tabela 2.6:</b> <i>Primers</i> utilizados para as mutações pontuais na PP2Ac $\beta$ e C-terminal da PP2Ac.....	54
<b>Tabela 3.1:</b> Descrição dos conjuntos de cristais coletados para a fosfatase PTP1B.....	63
<b>Tabela 3.2:</b> Condições de cristalização para a proteína TIPRL $\Delta$ N obtidas no ROBO LAB, nas quais houve sinais de cristalização nas concentrações entre 70 a 100 mg.mL <sup>-1</sup> de proteína. ....	71
<b>Tabela 3.3:</b> Estatísticas de coleta de dados para o conjunto a 2,66 Å.....	81
<b>Tabela 3.4:</b> Estatísticas de coleta e redução de dados para o conjunto a 2,15 Å.....	83
<b>Tabela 3.5:</b> Vínculos de <i>cross-linking</i> -MS para as amostras de TIPRL.....	91

## LISTA DE ABREVIATURAS

ASU: unidade assimétrica

DMSO: dimetilsulfóxido

DTT: ditioneitol

F<sub>calc</sub>: fator de estrutura calculado

F<sub>obs</sub>: fator de estrutura observado

HDX: Troca de hidrogênio por deutério

Hepes: ácido 4-(2-hidroxietil)piperazina-1-etano-sulfônico

HPV: vírus do papiloma humano

INCA: Instituto Nacional do Câncer

IPTG: isopropil-β-D-1-tiogalactopiranosídeo

LB: meio Lysogeny Broth, também referido como Luria Broth, Lennox Broth ou Luria-Bertani

MKK7: *Dual specificity mitogen-activated protein kinase kinase 7*

PDB: *Protein Data Bank*

PEG: polietilenoglicol

PMSF: fluoreto de fenil-metil-sulfonila

pNP: para-nitrofenol

pNPP: para-nitrofenil-fosfato

PP1: do inglês, *protein phosphatase 1* – proteína fosfatase 1

PP2A: do inglês, *protein phosphatase 2A* – proteína fosfatase 2A

PP2Ac: Subunidade catalítica da PP2A

PP4: do inglês, *protein phosphatase 4* – proteína fosfatase 4

PP4c: Subunidade catalítica da PP4

PP6: do inglês, *protein phosphatase 6* – proteína fosfatase 6

PP26: Subunidade catalítica da PP6

PTP: do inglês, *protein tyrosine phosphatase* – proteína tirosina fosfatase

PTP1B: proteína tirosina fosfatase 1B

R<sub>free</sub>: fator residual livre

RMSD: do inglês, *root mean square deviation* – desvio da raiz quadrada média

SPG: Tampão de fosfato, glicina e succinato

TEV: do inglês, *Tobacco Etch Virus*

TIPRL: do inglês, *TOR signaling pathway regulator-like*

Tris: tris(hidroximetil)aminometano

v/v: volume/volume

## Sumário

RESUMO.....	8
ABSTRACT .....	9
LISTA DE ILUSTRAÇÕES .....	10
LISTA DE TABELAS .....	12
LISTA DE ABREVIATURAS.....	13
CAPÍTULO 1 .....	18
1. Introdução.....	19
1.1 Fosforilação reversível .....	19
1.2 Câncer.....	20
1.3 Proteína Fosfatase 2A.....	22
1.3.1 Biogênese da PP2A.....	26
1.3.2 Proteína $\alpha 4$ .....	27
1.4 Proteína TIPRL .....	28
1.5 Proteína Tirosina Fosfatase PTP1B.....	29
1.6 Cristalografia de proteínas .....	33
CAPÍTULO 2 .....	37
Metodologia .....	37
2. Metodologia .....	38
2.1 Clones .....	38
2.2 Expressão de proteínas em sistema recombinante .....	40
2.2.1 Expressão em meio LB líquido .....	40
2.2.2 Expressão de TIPRL $\Delta$ N marcada com Selenometionina .....	41
2.3 Purificação das proteínas.....	41
2.4 Análises de proteínas por CD e DLS .....	44
2.5 Ensaio enzimático e Triagem de compostos para a proteína PTP1B... 45	45
2.6 Cristalografia de Proteínas.....	46
2.6.1 Cristalização .....	46
2.6.2 Coleta de dados e obtenção da estrutura tridimensional.....	47
2.7 SAXS.....	49

2.7.1	Coleta de dados.....	49
2.7.2	Análise das curvas.....	50
2.8	<i>Cross-linking</i> acoplado a MS.....	50
2.8.1	Reação de ligação cruzada utilizando DSS.....	50
2.9	Troca de Hidrogênio-Deutério HDX Análise por espectrometria de Massas .....	51
2.10	Mutagênese Sítio Dirigida .....	53
2.11	<i>GST-Pulldown</i> .....	54
CAPÍTULO 3 .....		56
3.	Resultados.....	57
3.1	Estudos com a proteína PTP1B .....	57
3.1.1	Ensaio de Inibição Enzimática .....	57
3.1.2	Estudos cristalográficos .....	60
3.2	Estudos com os Complexos de PP2Ac.....	67
3.2.1	Expressão Purificação e Cristalização.....	67
3.3	Estudos com a proteína TIPRL .....	70
3.3.1	Expressão, purificação e cristalização .....	70
3.3.2	Análises do estado oligomérico por SAXS e Gel Filtração .....	76
3.3.3	Coleta de dados.....	80
3.3.4	Estrutura .....	84
3.3.5	<i>Cross-linking</i> acoplado a MS .....	89
3.3.6	Troca de Hidrogênio por Deutério.....	91
3.3.7	Mutagênese sítio dirigida associada a <i>GST-pulldown</i> .....	97
CAPÍTULO 4 .....		101
4.	Manuscrito TIPRL.....	102
CAPÍTULO 5 .....		132
5.	Discussão .....	133
5.1	Discussão dos resultados dos estudos com PTP1B .....	133
5.2	Discussão dos resultados dos estudos com TIPRL .....	135
5.2.1	Dimerização.....	135
5.2.2	Obtenção da Estrutura.....	136
5.2.3	Estrutura .....	138
5.2.4	Provável função biológica da TIPRL.....	140
CAPÍTULO 6 .....		144



6. Conclusões Perspectivas.....	145
7. Referências.....	147

# **CAPÍTULO 1**

## **Introdução**

# 1. Introdução

## 1.1 Fosforilação reversível

O conceito de fosforilação de proteínas, descoberto por Edmond Fischer e Edwin Krebs na década de 1950, surgiu a partir do ensaio *in vitro* com ATP e uma "enzima conversora" (nomeada posteriormente fosforilase quinase) que convertia a fosforilase *b* em fosforilase *a*. Nesta reação houve a transferência de um grupo fosfato do ATP para a fosforilase *b*, resultando na fosforilase *a*, uma fosfoproteína. Curiosamente, a enzima que converte fosforilase *a* de volta para *b*, chamada "*PR enzyme*" (fosforilase fosfatase), tinha sido relatada uma década antes, em 1943, embora a natureza química da reação de fosforilação tenha-se mantido enigmática até a verificação de que o fosfato inorgânico era um produto da reação (Shi 2009).

A fosforilação reversível das proteínas desempenha um papel crucial na regulação da sinalização celular. Ela é responsável pela maioria das transduções de sinal em células eucarióticas, controla processos celulares como metabolismo, transcrição, progressão do ciclo celular, rearranjo do citoesqueleto e mobilidade celular, apoptose e diferenciação. A fosforilação de proteínas também desempenha um papel crítico na comunicação intercelular durante o desenvolvimento, nas respostas fisiológicas, na homeostase, no funcionamento dos sistemas nervoso e imunológico (Hunter 1995; Manning *et al.* 2002). Em células normais, fosforregulação é rigorosamente controlada por uma rede de proteínas quinases contrabalançadas por várias proteínas fosfatases (Haesen *et al.* 2014). Mutações e desregulações nas proteínas quinases e fosfatases desempenham um papel fundamental no desenvolvimento de doenças humanas, tais como o câncer e a diabetes. Devido a isso, os estudos das proteínas responsáveis pela fosforilação reversível possibilitam o desenvolvimento de novos fármacos que podem ativar ou inibir vias de sinalização e auxiliar no tratamento de doenças (Manning *et al.* 2002). A importância dos estudos da fosforilação rendeu a Krebs e Fischer o prêmio Nobel de fisiologia e medicina em 1992 (Salton 2005).

A fosforilação de proteínas eucarióticas ocorre nos resíduos de aminoácidos serina, treonina e tirosina, na seguinte proporção, respectivamente

79,3%, 16,9% e 3,8% (Olsen *et al.* 2006). Aproximadamente, 17.000 proteínas têm pelo menos um resíduo de aminoácido que pode ser fosforilado, sendo assim possível substrato para as proteínas quinases e fosfatases (Hornbeck *et al.* 2012). As proteínas quinases formam uma das maiores famílias de genes em eucariotos, responsáveis por cerca de 2% de todo o genoma: são expressas 518 proteínas quinases no genoma humano; deste total, 428 fosforilam resíduos de serina e treonina e as outras 90 resíduos de tirosina (Manning *et al.* 2002). Em contraste, no genoma humano existem 179 genes para as proteínas fosfatases (Sacco *et al.* 2012).

As proteínas fosfatases se dividem em quatro famílias de acordo com seus domínios catalíticos, identidade sequencial e substratos. As Proteínas Fosfatases PPP (*Phospho Protein Phosphatase*) e PP2C, as quais removem o fosfato dos resíduos de serina e treonina; as PTPs (*Phospho Tyrosine Phosphatase*) que desfosforilam resíduos de tirosina; e a família NIF (*NLI interacting fator-like phosphatase*), que possui um domínio de fosfatase e age na RNA polimerase II. Esta tese apresenta estudos referentes a duas proteínas fosfatases, a Proteína Fosfatase PP2A, um importante supressor tumoral, e a Proteína Tirosina Fosfatase PTP1B, a qual é superexpressa em alguns tipos de tumores.

## 1.2 Câncer

A descoberta, no século XIX, de que todas as células de um organismo pluricelular descendem de um óvulo fecundado levaram à conclusão de que tumores não eram corpos estranhos, mas sim um crescimento derivado de tecidos normais. As anomalias celulares de tumores apontaram o câncer como sendo uma doença da célula em mau funcionamento (Weinberg 2008). O câncer pode ser considerado como uma classe de doenças do DNA e do metabolismo celular (Junqueira e Carneiro 2005). Ele se forma a partir de uma célula que sofreu mutação, proliferou, e suas descendentes foram acumulando mutações até que apareceram novas células que não obedecem aos mecanismos de controle do ciclo celular e se multiplicam continuamente (Junqueira e Carneiro 2005). Há mais de cem diferentes tipos de câncer caracterizados. A nomenclatura desses diferentes tipos é dividida em quatro

grandes grupos de acordo com a origem embrionária do tecido, epitelial, mesenquimal, hematopoiética e neuroectodermal. Todos os tipos celulares do organismo podem originar tumores malignos, todavia em humanos o mais comum, por volta de 80% de todos os tipos de câncer, são os tipos celulares de origem epitelial, devido à exposição ao meio e à contínua divisão para repor as células mortas (Weinberg 2008).

Os tumores podem ser tanto benignos (localizados e não invasivos) como malignos (invasivos e metastáticos). A metástase que é a capacidade de invadir tecidos e órgãos vizinhos ou distantes, semeada pelos tumores malignos responsável por quase todos os óbitos causados pela doença (Weinberg 2008; Junqueira e Carneiro 2005).

As células responsáveis pelo câncer podem ser caracterizadas por dez *hallmarks*, marcas biológicas, que são capacidades biológicas adquiridas durante o desenvolvimento do tumor. As células cancerosas não respondem à sinalização dos supressores de crescimento, têm sinalização proliferativa, evitam ser destruídas pelo sistema imunológico, são imortais, causam inflamações, invadem outros tecidos provocando a metástase, induzem a angiogênese, possuem genomas instáveis e com mutações, resistem à morte celular e o metabolismo energético é desregulado (Hanahan e Weinberg 2000, 2011). Conhecer bem a doença é fundamental para o desenvolvimento de novos tratamentos.

A biologia do câncer não é simples, envolvendo diversas vias metabólicas e mecanismos de sinalização. Um deles é a fosforilação, a qual é regulada pelas proteínas quinases e fosfatases. Essas duas classes de enzimas têm representantes que podem atuar como supressores ou ativadores tumorais. Portanto, são importantes alvos no desenvolvimento de novas drogas para o tratamento dessa patologia. Terapias baseadas em inibidores de quinases têm sido extremamente bem sucedidas no tratamento da leucemia mielóide crônica (LMC) (Jabbour e Kantarjian 2014); como as fosfatases têm papel antagônico na sinalização, também são alvo de estudos para terapias baseadas na inibição da catálise. No entanto, vários desafios permanecem no desenvolvimento de novas terapias, incluindo o desenvolvimento de resistência à droga, a falta de seletividade

do inibidor ou de eficácia e dificuldade de validação de alvos de drogas (Kamb 2007).

A incidência de alguns cânceres pediátricos é distribuída igualmente na população mundial devido à origem genética. Já a incidência dos tumores malignos em adultos varia drasticamente pelos países, pois a origem dos cânceres tem fatores genéticos e ambientais, sendo este último o mais determinante (Weinberg 2008). O estilo de vida pode ser determinante em alguns casos de tumores, como o câncer primário de bexiga que é causado em sua maioria pelo consumo de cigarros. Os fatores ambientais que podem desencadear a doença são vírus, como o HPV (papilomavírus humano), radiação ionizante, como os Raios X, químicos como as substâncias presentes em cigarros. O Brasil por ser um país tropical com grande incidência dos raios solares, a questão cultural da pele bronzeada ser considerada como bela e o alto custo de protetores solares, fazem com que o câncer de pele não melanoma seja o de maior incidência no país (INCA 2015).

### 1.3 Proteína Fosfatase 2A

A Família das Fosfoproteínas Fosfatases – PPP (*Phospho Protein Phosphatase*) está presente em todos os seres eucariotos e também foi identificada no genoma de *Archaea* (Pereira *et. al.*, 2011; Kennelly 2003). Essa família é definida por três motivos característicos em sua sequência peptídica, **GDxHG(x)<sub>23</sub>GDxVDRG(x)<sub>25</sub>GNHE**. Os resíduos de aminoácidos em negrito estão envolvidos na coordenação com metal, um cátion bivalente (Goldberg *et. al.*, 1995; Xu *et. al.*, 2006; Shi 2009; Pereira *et. al.*, 2011). As enzimas desse grupo desfosforilam resíduos de serina/treonina de substratos proteicos, regulam várias vias metabólicas importantes e são inibidas por ácido ocadaíco (Shi 2009). Além disso, o extremo C-terminal é conservado, formado pelos resíduos de aminoácidos TPDYFL. As fosfatases PP2A, PP4, PP6 e PP1 são as representantes dessa família.

A fosforilação na célula normal ocorre em grande maioria, 96%, nos resíduos de aminoácidos serina e/ou treonina, este processo é coordenado pelas enzimas Ser/Thr quinases e Ser/Thr fosfatases. Paradoxalmente, é estimado que o

número de genes humanos de Ser/Thr fosfatases sejam por volta de 40 enquanto o de quinases dessa classe sejam de 428, o que leva a pensar que as fosfatases são promíscuas e não possuem especificidade de substrato. Porém a família PP2A representa um bom exemplo de como este paradoxo pode ser explicado, uma vez que estas fosfatases adquiriram especificidade através da montagem de um grande número de diferentes holoenzimas de múltiplas subunidades a partir de um número limitado de subunidades catalíticas.

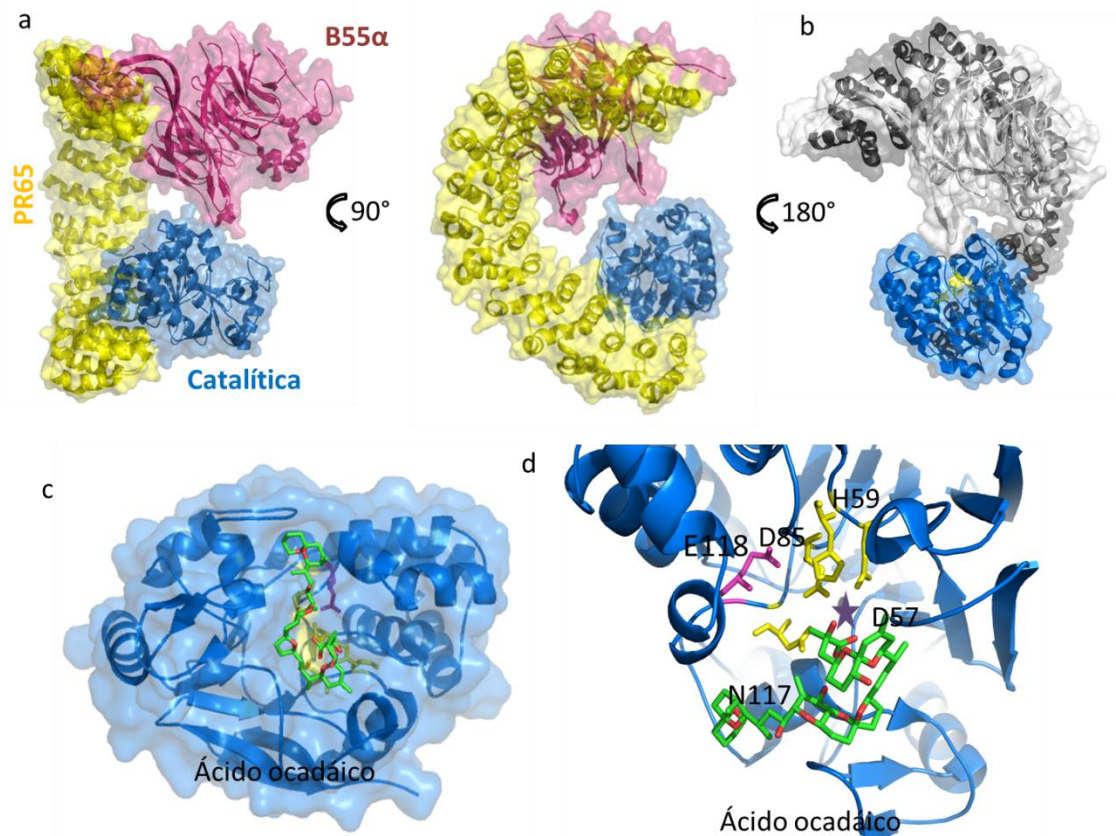
A Proteína Fosfatase 2A (PP2A) é a mais importante representante da família Proteína Fosfatase e é responsável pela regulação de várias vias de sinalização. Por exemplo, algumas de suas proteínas substratos são a p53 (importante supressor tumoral, regula as vias de crescimento celular e apoptose), Akt (quinase envolvida em diversas vias de sinalização celular como as vias de apoptose e estresse oxidativo) e c-Myc (Fator de transcrição do DNA e ativador da transcrição de genes relacionados ao crescimento celular) (Seshacharyulu *et. al.*, 2013). Além disso, a PP2A desempenha diversas funções relacionadas à proliferação, mobilidade, controle do ciclo e morte celular, dinâmica do citoesqueleto e também é importante supressor tumoral (Janssens e Goris 2001; Janssens *et. al.*, 2005; Janssens e Rebollo 2012; Xing *et. al.*, 2006; Migueleti *et. al.*, 2012). A PP2A tem função pro-apoptótica, regulando a p53, proteína com funções de supressor tumoral, que ativa as vias de apoptose.

Além de suprimir tumores, a PP2A também está envolvida em algumas doenças humanas como o mal de Alzheimer e arritmias cardíacas. Na primeira patologia, a PP2A, subunidade regulatória B55 $\alpha$ , converte o excesso de fosforilação causado por uma quinase na proteína TAU. No Alzheimer, a PP2A não desfosforila a TAU isso faz com que a TAU fosforilada em excesso forme fibras amiloides no cérebro ocasionando a morte de neurônios (Torrent e Ferrer 2012). Já nas arritmias cardíacas, a superexpressão do microRNA, *miR-1*, no músculo cardíaco, reduz os níveis de PP2A, subunidade regulatória B56 $\alpha$ , o que gera um aumento na fosforilação dos receptores de rianodina nos canais de liberação de cálcio causando um desequilíbrio no fluxo de Ca<sup>2+</sup>, o qual leva às arritmias cardíacas (Terentyev *et. al.*, 2009).

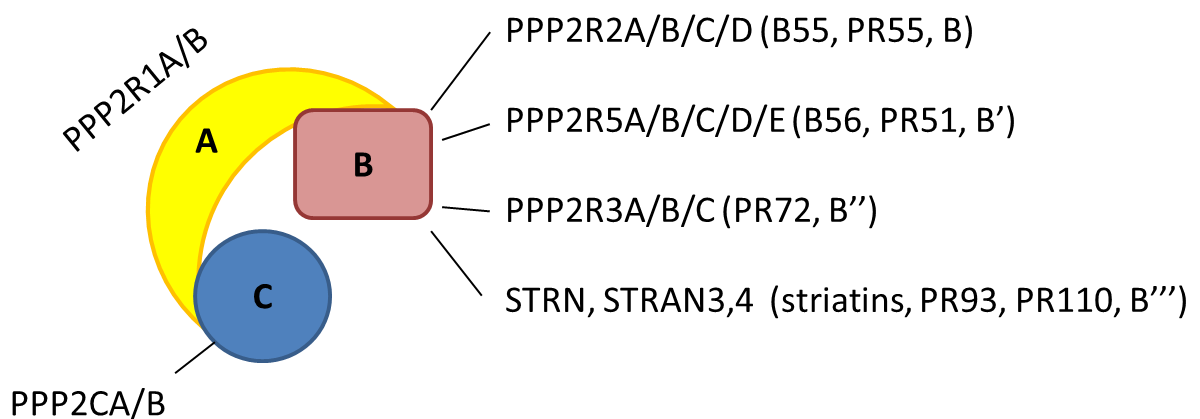
Estruturalmente, a PP2A forma em torno de 96 diferentes complexos proteicos com outras proteínas, podendo ter atividade catalítica ou não (Sents *et.*

*al.*, 2012) O heterotrímero funcional da PP2A, A:B:C, é formado pelas subunidades A - ancoragem, B – regulatória e C – catalítica (Virshup 2000) (Figura 1.1). Em células humanas, as subunidades B são codificadas por 15 genes diferentes, os quais dão origem a 46 isoformas diferentes através da utilização de promotores alternativos de genes e eventos de *splicing* alternativo (Lambrecht *et. al.*, 2013). Com base na homologia de sequências, são divididas em quatro famílias distintas, chamadas B (ou B55, ou PR55, ou pelo nome do gene: PPP2R2), B' (ou B56, ou PR61, ou pelo nome do gene: PPP2R5), B'' (ou PR72, ou pelo nome do gene: PPP2R3) e B''' (ou *striatins*, STRN). As subunidades do tipo B são subunidades reguladoras, no sentido de que ditam a especificidade do substrato da subunidade C da PP2A. Elas são, geralmente, expressas em células específicas e podem ser encontradas em diferentes estruturas subcelulares (citoplasma, núcleo, membrana plasmática, mitocôndrias, complexo de Golgi, retículo endoplasmático e citoesqueleto). As subunidades A são codificadas por dois genes diferentes, PPP2R1A e PPP2R1B, dando origem a duas isoformas A $\alpha$  e A $\beta$ . Além disso, as subunidades C catalíticas são codificadas por dois genes, PPP2CA e PPP2CB, dando origem a 2 isoformas distintas C $\alpha$ 1 e C $\beta$  (Figura 1.2 e Tabela 1.1).





**Figura 1.1:** Heterotrímtero funcional da PP2A. (a) Heterotrímtero funcional com as subunidades A, em amarelo, B em rosa e a catalítica c, em azul, adaptado de Xu 2008 (código PDB: 3DW8). (b) Destaque para a porção catalítica em azul e a entrada do sítio ativo em amarelo. (c) Sítio catalítico da PP2Ac em amarelo ocupado pelo ácido ocadaico. (d) Zoom do sítio catalítico. Em amarelo são destacados os resíduos D57, H59, D85 e N117 que coordenam com o metal bivalente representado por uma estrela roxa, o resíduo catalítico em magenta E118 e o ácido ocadaico em verde. Adaptado de Xu 2008 (código PDB: 2IE4).



**Figura 1.2:** Combinações para a formação do heterotrímtero funcional da PP2A. Adaptado de Sents et. al., 2012.

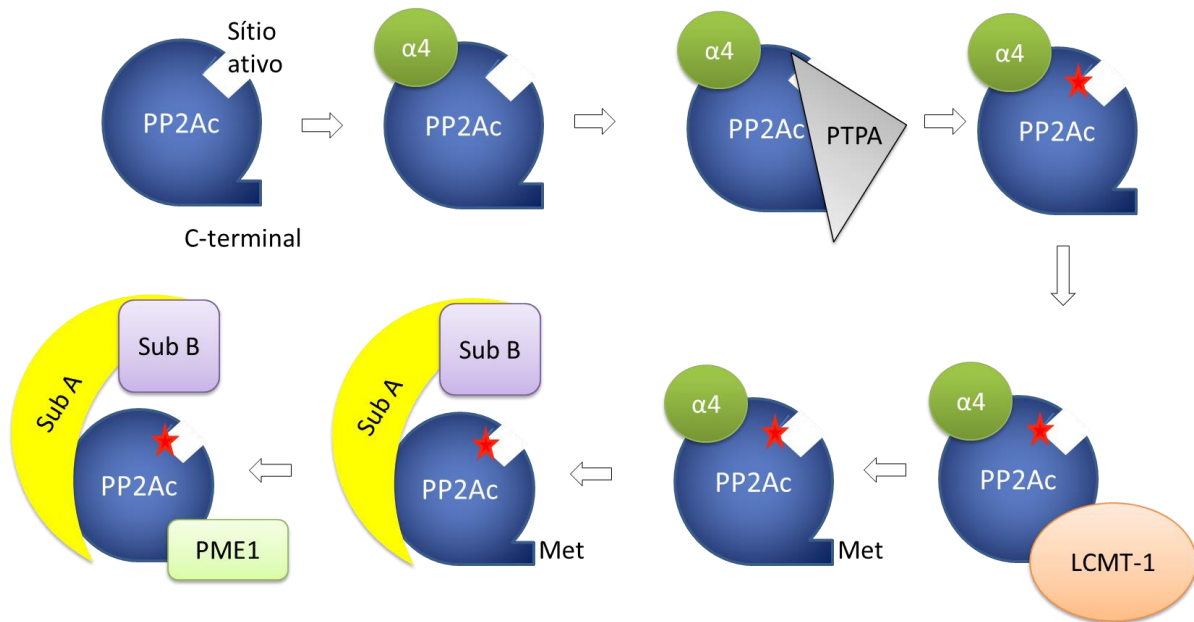
**Tabela 1.1:** Lista de genes que codificam as isoformas das subunidades ABC para montagem do heterotrímero funcional da PP2A.

Subunidade	Gene	Proteína	<i>Splicing</i> alternativo
Ancoragem A	PPP2R1A	A $\alpha$	
	PPP2R1B	A $\beta$	
Regulatória B	PPP2R2A/B/C/D	B ou B55	$\alpha$ (2); $\beta$ (7); $\gamma$ (4); $\delta$ (1)
	PPP2R5A/B/C/D/E	B' ou B56	$\alpha$ (2); $\beta$ (2); $\gamma$ (5); $\delta$ (3); $\epsilon$ (3)
	PPP2R3A/B/C	B'' ou PR72	$\alpha$ (3); $\beta$ (3); $\gamma$ (2)
	STRN	Striatin	
	STRN3	Striatin-3	$\alpha$ ; $\beta$
	STRN4	Striatin-4	1; 2; 3
Catalítica C	PPP2CA	C $\alpha$	
	PPP2CB	C $\beta$	

### 1.3.1 Biogênese da PP2A

A montagem da holoenzima PP2A A:B:C ativa é um processo fortemente regulado, o qual requer uma série de interações proteína-proteína e modificações pós-traducionais na subunidade catalítica, no entanto esse processo ainda não está totalmente compreendido. O primeiro passo é a sua interação com a proteína  $\alpha 4$  (Figura 1.3), que estabiliza uma conformação inativa da fosfatase e a protege da degradação proteossômica até a montagem do complexo heterotrímero funcional (Kong *et. al.*, 2009; Jiang *et. al.*, 2013). As interações que se seguem são com PTPA (*phosphotyrosyl phosphatase activator*), a qual promove a adição do cátion bivalente  $Mg^{2+}$  e LCMT-1 (*leucyl carboxymethyl transferase*), que adiciona o grupo metil ao extremo C-terminal (leucina 309).

A proteína PME-1 (*Protein Phosphatase type 2A methylesterase 1*) remove a metilação promovida pela LCMT-1, dependendo de qual subunidade regulatória está ligada à catalítica. Outra modificação pós-traducional da subunidade catalítica é a fosforilação da treonina 304 e da tirosina 307, possivelmente catalisadas pelas quinases CDK e RTK, que afetam a preferência por subunidades regulatórias específicas. Ao fim desse processo, são incorporadas as subunidades regulatória e de ancoragem e as outras proteínas envolvidas ( $\alpha 4$ , PTPA, LCMT-1) são dissociadas e possivelmente recicladas para a formação de novos complexos (Sents *et. al.*, 2012).



**Figura 1.3:** Esquema da biogênese do heterotrímero funcional da fosfatase PP2A. Não é sabido em que etapa a  $\alpha 4$  é dissociada da PP2Ac. A estrela em vermelho representa o cátion bivalente.

### 1.3.2 Proteína $\alpha 4$

A proteína  $\alpha 4$  (também conhecida por IGBP1 *Immunoglobulin-alpha-binding protein 1* em mamíferos e Tap42 – *Phosphatase 2A-associated protein of 42kDa* em fungos) é constituída de uma porção N-terminal em TPR (tetratricopeptídeo), composta por sete  $\alpha$ -hélices antiparalelas (Yang *et. al.*, 2007) que interagem com a PP2Ac através de resíduos carregados de aminoácidos, e uma outra região intrinsicamente desenovelada no C-terminal (Smetana *et. al.*, 2006). A estrutura da  $\alpha 4$  de levedura foi determinada em 2007 e a de rato em 2011 (códigos PDB 2V0P e 3QC1, respectivamente), por técnicas cristalográficas (Yang *et. al.*, 2007). Ambas referem-se à porção com enovelamento estável composta pelos resíduos de aminoácidos 1 a 222, denominada  $\alpha 4\Delta 222$  (Smetana *et. al.*, 2006). A  $\alpha 4$  forma um complexo estável com PP2Ac independente das subunidades A e B (Sents *et. al.*, 2012). Além disso, auxilia no correto enovelamento da PP2Ac e desempenha papel importante para assegurar a estabilidade estrutural da PP2Ac no ambiente celular (Jiang *et. al.*, 2013).

A superexpressão de  $\alpha 4$  já foi observada em alguns tipos de câncer como carcinomas hepatocelular, de pulmão e de mama. Em contrapartida, o *knockout* do

gene da  $\alpha 4$  em levedura e camundongo é letal (Sents *et. al.*, 2012), mostrando sua importância no metabolismo celular.

## 1.4 Proteína TIPRL

A proteína TIPRL (*target of rapamycin signaling pathway regulator-like*) é um regulador negativo da PP2Ac. A subunidade catalítica interage com a proteína TIPRL, formando um complexo sem atividade, o qual também se liga à  $\alpha 4$ . A PP2Ac interage com ambas as proteínas, mas as proteínas  $\alpha 4$  e TIPRL não formam um complexo na ausência da PP2Ac como ocorre nos ortólogos de levedura TAP42 e TIP41 respectivamente (Smetana e Zanchin 2007; McConnell *et. al.*, 2007; Jacinto *et. al.*, 2001). O complexo inativo entre as proteínas PP2Ac e TIPRL ocorre, preferencialmente, quando a subunidade catalítica da PP2Ac está metilada e, somente, entre a subunidade C e a TIPRL, não há formação do complexo quando as subunidades A e/ou B estão ligadas a C (Nakashima *et. al.*, al. 2013).

Smetana e Zanchin (Smetana e Zanchin 2007) sugerem que a região dos resíduos de aminoácidos 210-309, correspondente ao C-terminal da PP2Ac, seja fundamental para a interação com a TIPRL, já que essa região é a menor porção que interage com a TIPRL dentre as construções testadas no trabalho. Além disso, acredita-se que o C-terminal da TIPRL seja importante para a formação do complexo PP2Ac:TIPRL, pois a TIPRL de 20 kDa, TIP\_i2, não forma complexo com a PP2Ac (McConnell *et. al.*, 2007). Essa proteína TIP\_i2 é uma variante de *splicing*, que é transcrita até o resíduo de aminoácido 178, portanto não há, em sua cadeia peptídica, os 94 resíduos de aminoácidos correspondentes à porção C-terminal da TIPRL. Assim, acredita-se que a interface de interação do complexo PP2Ac:TIPRL seja através das porções C-terminais de cada proteína. Essa ideia é reforçada pelo fato da proteína TIPRL formar complexos heterodiméricos com PP4 e PP6, as quais possuem a mesma sequência de aminoácidos na porção C-terminal que a PP2Ac, e por não formar complexo com a PP1, que tem o seu C-terminal diferente das demais fosfatases da família (McConnell *et. al.*, 2007; Smetana e Zanchin 2007).

A estrutura tridimensional da proteína PP2A subunidade catalítica C complexada com a subunidade de ancoragem A já foi elucidada em 2006 por Xing e colaboradores (Xing *et. al.*, 2006) (código PDB 2IE4). Outras estruturas de

complexos entre a subunidade catalítica e diferentes proteínas também já foram resolvidas, como as seguintes: PME-1 (código PDB 3C5W) em 2008 (Xing *et. al.*, 2008), subunidade regulatória B55 (código PDB 3DW8) também em 2008 (Xu *et. al.*, 2008), LMCT-1 (código PDB 3P71) em 2001 (Stanevich *et. al.*, 2011),  $\alpha 4$  (código PDB 4IYP) em 2013 (Jiang *et. al.*, 2013) e mais recentemente, 2014, PTPA (código PDB 4LAC) (Guo *et. al.*, 2014). Entretanto, a estrutura do complexo heterodimérico PP2Ac:TIPRL ainda não foi elucidada.

O papel biológico da proteína TIPRL ainda não foi esclarecido. Acredita-se que a TIPRL seja uma proteína com características antiapoptose, pois em células de hepatocarcinoma, as quais apresentaram superexpressão de TIPRL, não houve morte celular na presença de TRAIL (*Tumor Necrosis Factor (TNF)-related apoptosis-inducing ligand*). Além disso, o *knockout* do gene que codifica a TIPRL em embriões de *zebrafish* resultou em um aumento de apoptose (Song *et. al.*, 2012). Todavia, vários aspectos dela ainda são desconhecidos. Sents e colaboradores, 2012, (Sents *et. al.*, 2012) sugerem que a TIPRL tenha um papel na biogênese da PP2Ac ou na degradação depois da desmontagem da holoenzima, mas ainda não está totalmente esclarecido o seu papel fisiológico. Esta tese apresenta novos estudos sobre a proteína TIPRL, como sua estrutura tridimensional inédita e suas interações com a PP2Ac. Essas descobertas novas colaboram para o entendimento de função biológica da TIPRL.

## 1.5 Proteína Tirosina Fosfatase PTP1B

A família das proteínas tirosina fosfatases (PTPs) reúne as enzimas responsáveis pela remoção do grupo fosfato de resíduos de tirosina em proteínas. O sítio catalítico dessa família é altamente conservado por uma sequência consenso composta pelos resíduos [I/V]HCXXGXXR[S/T], os resíduos em negrito são os catalíticos, cisteína e arginina (Tonks 2013). Esta sequência forma o *loop* catalítico, responsável pela remoção do grupo fosfato do resíduo de tirosina da proteína substrato. Essa família é formada por 107 proteínas (Ostman *et. al.*, 2006), sendo a proteína PTP1B, a representante desse grupo nesta tese.

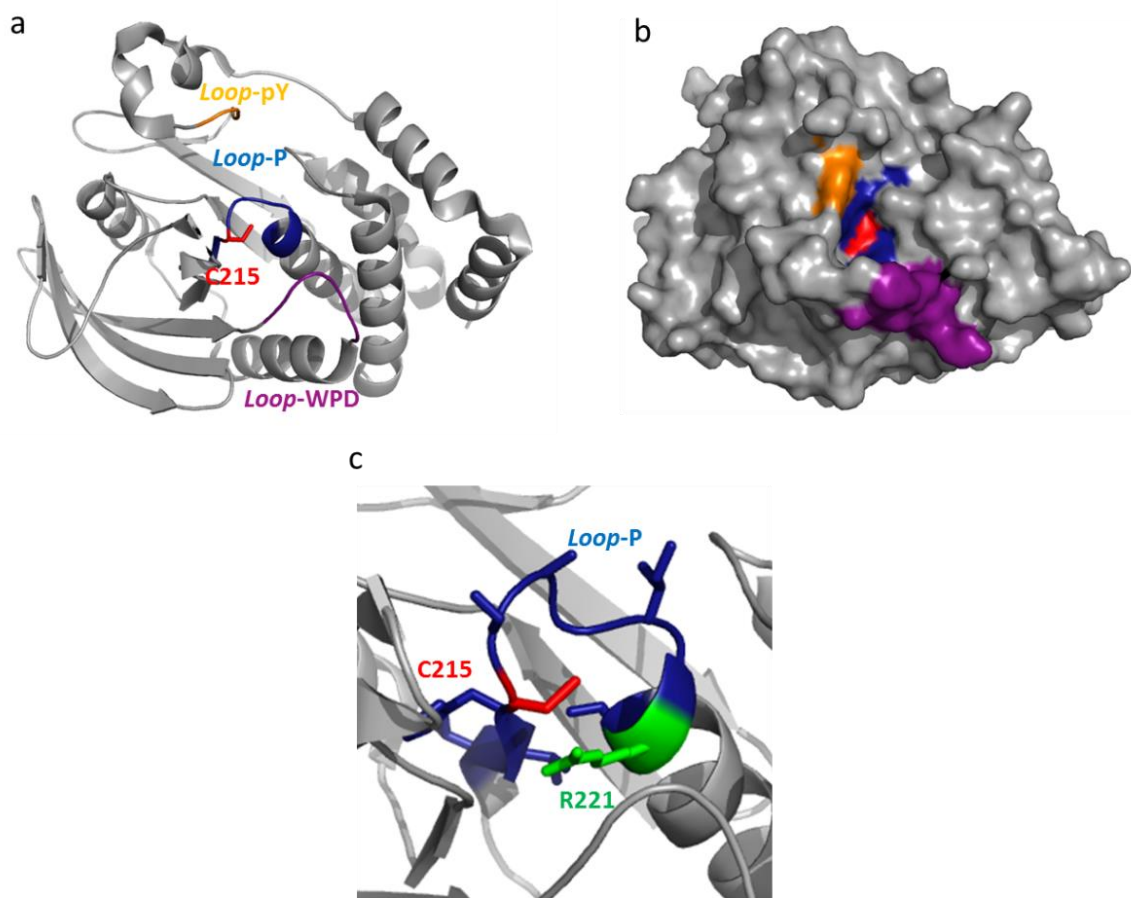
A fosfatase PTP1B, também denominada PTPN1, é expressa pelo gene PTPN1(PTP1B) em várias células eucariotas. Ela é uma proteína citoplasmática de 50 kDa, a qual fica ancorada no retículo endoplasmático (RE) por meio de uma cauda hidrofóbica com cerca de 150 resíduos de aminoácidos que compõem a região do C-terminal. A extensa rede RE celular e a orientação topológica do sítio catalítico da PTP1B voltado para o citossol orientam sua atividade enzimática para uma variedade de substratos, dentre os quais a PP2Ac, que tem seu resíduo de aminoácido Y307 desfosforilado por essa tirosina fosfatase (Wong *et. al.*, 2010). Evidências recentes sugerem que a PTP1B tenha a função de adesão em locais de contato através da associação do RE com microtúbulos. Na maioria dos casos, os substratos de PTP1B foram identificados através de mutação sítio-dirigida no resíduo D181A, a qual estabiliza um complexo enzima:substrato. Dentre as proteínas substrato identificadas estão os fatores de crescimento epidermal e receptores de fatores de crescimento derivados de plaquetas, receptores de insulina e complexos de adesão de caderina (Monteleone *et. al.*, 2012).

A PTP1B está relacionada a algumas doenças humanas, tais como resistência a insulina, diabetes e câncer. O *knockout* do gene que codifica a PTP1B em camundongos gerou indivíduos saudáveis que apresentaram aumento da sensibilidade à insulina e resistência à obesidade induzidos por uma dieta rica em gordura. Isto sugere que a inibição da PTP1B pode promover a sinalização em estados de resistência à insulina e oferecer uma nova abordagem para o tratamento de diabetes e obesidade (Elchebly *et. al.*, 1999; Klamann *et. al.*, 2000). A PTP1B é superexpressa em tumores de mama relacionados com a expressão de HER2 (quinase associada com o aumento de síntese de proteínas e crescimento celular), que são tumores agressivos. Além disso, o gene PTP1B está localizado no cromossomo 20q13, uma região frequentemente amplificada em câncer de mama e está associado a um prognóstico pouco animador (Tonks 2013). Camundongos que expressam alelos ativados de HER2 em glândulas mamárias desenvolvem múltiplos tumores mamários e metástases para o pulmão. No entanto, quando tais animais foram cruzados com camundongos *knockout* para PTP1B, o desenvolvimento do tumor foi retardado e a incidência de metástases no pulmão foi reduzida (Julien *et. al.*, 2007; Bentires-Alj e Neel 2007). Portanto, é esperado que, ao inibir a função da PTP1B com fármacos, a sinalização em condições de resistência à insulina melhore

e atenua a malignidade da tumorigênese mamária em câncer de mama positivo para HER2 (Tonks 2013).

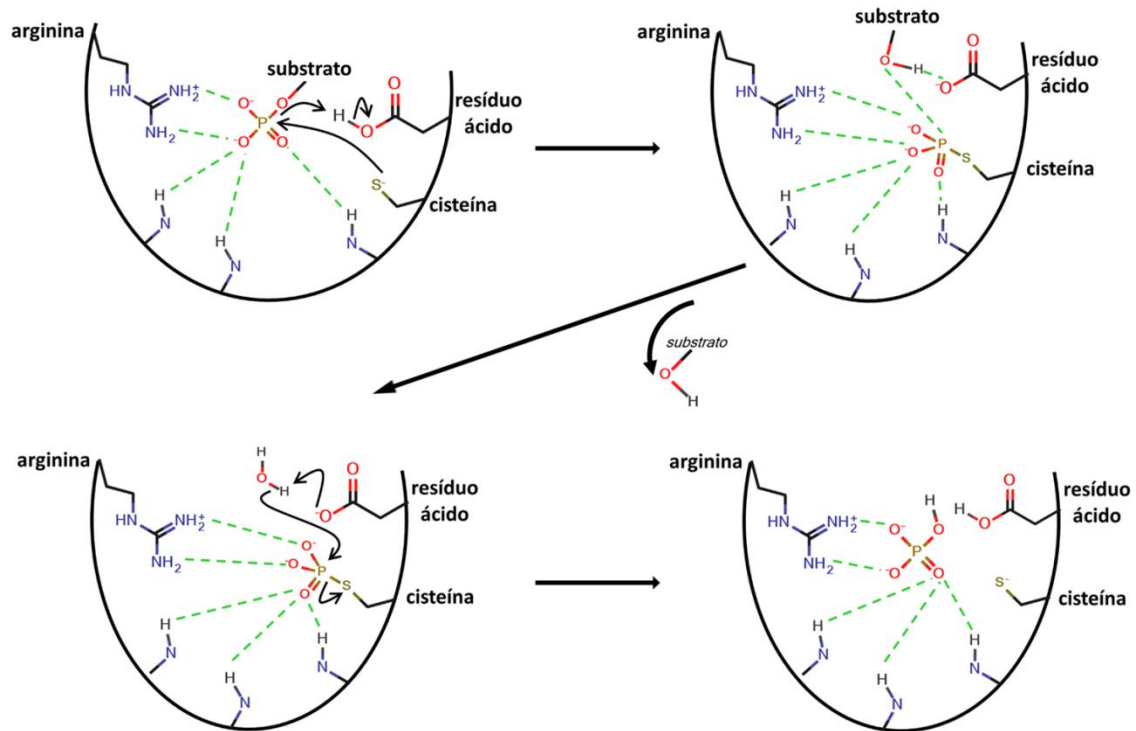
Entre as proteínas fosfatases que compõem a família das PTPs, PTP1B é a proteína mais estudada do ponto de vista estrutural. Os elementos de estrutura secundária chave desta proteína são mostrados na Figura 1.4. Dentre estes elementos estruturais, destacam-se o *loop*-P (em azul e em bastão vermelho a cisteína catalítica), o *loop* de reconhecimento do substrato (*pY recognition loop* – em laranja) e o *loop*-WDP (em roxo). O *loop*-P catalítico é formado pelos resíduos<sup>213</sup> VHCSAGIGRS<sup>222</sup>, reconhece o fosfato do substrato e o resíduo de cisteína 215, quando desprotonado, promove o ataque nucleofílico ao fósforo da proteína fosforilada. Em seguida ocorre hidrólise do intermediário cistenil-fosfato por adição de uma molécula de água, completando o ciclo catalítico pela regeneração da enzima ativa (Tonks 2013) Figura 1.5. Na presença de agentes oxidantes, o resíduo de cisteína é oxidado e a enzima PTP1B tende a perder sua atividade (van Montfort *et. al.*, 2003). A especificidade da PTP1B à tirosina fosforilada do substrato é dada pelo resíduo de aminoácido tirosina 46, o qual faz parte do *loop*-pY, este resíduo determina a profundidade do sítio catalítico em 9 Å (Zhang 2002), deste modo, os resíduos menores fosfoserina e fosfotreonina não conseguem chegar até a cavidade catalítica para sofrer a desfosforilação (Tonks 2013).

A primeira estrutura tridimensional da proteína fosfatase PTP1B humana foi resolvida por Barford e Tonks, 1994, (Barford *et al.* 1994) através da técnica de difração de Raios X. Atualmente, há 128 estruturas de PTP1B humana depositadas no *Protein Data Bank* (PDB [www.rcsb.org](http://www.rcsb.org)), tanto em sua forma apo (sem ligantes) quanto na forma holo (em complexos com ligantes). Muitos desses ligantes são inibidores da PTP1B e já foram descritos, pertencentes a uma grande gama de classes químicas, incluindo produtos naturais e derivados (ex.: diterpenos e derivados (Han *et. al.*, 2005)), peptídeo-miméticos (Larsen *et. al.*, 2003; Liljebris *et. al.*, 2002; Larsen *et. al.*, 2002; Bleasdale *et. al.*, 2001), entre outros (Heneberg 2009; Lee e Wang 2007). Nesta tese, buscou-se encontrar novos ligantes para a proteína tirosina fosfatase PTP1B humana e estruturas cristalográficas proteína:ligante.



**Figura 1.4:** Estrutura da PTP-1B e elementos estruturais importantes para catálise e desenho de inibidores. (a) elementos de estrutura secundária da estrutura da PTP-1B. O carbono alfa da cisteína catalítica é mostrado como bastão em vermelho e o *loop-P* catalítico em azul, *loop-pY* de reconhecimento de substrato em laranja e o *loop-WPD*, que fecha o sítio durante a catálise, em roxo. (b) Superfície da PTP1B, C215 no fundo do sítio catalítico com profundidade de 9Å regulada pela Y46 em laranja. (c) Zoom do *loop* catalítico com a cisteína C215 em vermelho e a arginina R221 em verde (código PDB: 3A5J).





**Figura 1.5:** Esquema da reação de remoção de grupo fosfato pela fosfatase PTP1B. O sítio catalítico reconhece o fosfato do substrato e o resíduo de cisteína 215, quando desprotonado, promove o ataque nucleofílico ao fósforo da proteína fosforilada. Em seguida ocorre hidrólise do intermediário cistenil-fosfato por adição de uma molécula de água, completando o ciclo catalítico pela regeneração da enzima ativa (Fonseca, 2015).

## 1.6 Cristalografia de proteínas

A cristalografia é uma ciência relativamente recente. Os primeiros a determinaram estruturas cristalinas foram Willian Bragg e seu filho Lawrence Bragg, em 1914, com as resolução das estruturas das moléculas do cloreto de sódio (NaCl), sulfeto de zinco e do diamante (Ewald 1962). No ano seguinte, ganharam o Nobel de Física por esses estudos. Com o passar dos anos, a técnica evoluiu e permitiu resolver a estrutura de moléculas biológicas como o colesterol (1937), vitamina B12 (1945), penicilina (1954) e insulina (1969) pela química Dorothy Hodgkin (Glusker 1994), DNA (1953) por Francis Crick, James Watson e Rosalind Franklin (Watson e Crick 1953; Franklin e Gosling 1953) e de proteínas como a mioglobina (1957) por Max Perutz e John Kendrew (Kendrew *et. al.*, 1958).

Era extremamente difícil resolver uma estrutura de proteína devido às limitações técnicas. Hodgkin levou 35 anos para resolver a estrutura da insulina. Em 1953, Max Perutz descobriu que era possível solucionar o problema das fases em cristalografia de proteína pelo método de substituição isomorfa com átomos pesados. Esperava-se que os cristalógrafos estruturais resolvessem muitas estruturas de proteínas com essa técnica de sinal isomorfo, entretanto isso não aconteceu. Apenas quatorze estruturas de proteínas foram resolvidas até 1970 (Sjolin 1993).

A dificuldade em resolver a estrutura de proteínas foi parcialmente superada com o advento das técnicas de DNA recombinante, automação em processos de purificação e cristalização, síncrotrons com brilho mais intenso, detectores sensíveis, *softwares* para o tratamento matemático dos dados, computadores com processadores cada vez mais velozes para tratar os dados, bancos de estrutura como o PDB, que facilita a resolução de estrutura pelo método de substituição molecular. Atualmente, o número de novas estruturas de proteínas cresce de modo exponencial devido a essas facilidades, todavia sem perder a sua importância. Nos últimos anos, vários pesquisadores receberam o Nobel por seus trabalhos envolvendo cristalografia de proteínas por difração de Raios X, como o prêmio de 2009, dado em reconhecimento ao trabalho de Venkatraman Ramakrishnan, Thomas Steitz e Ada Yonath, por resolverem a estrutura do ribossomo bacteriano.

A importância de resolver uma estrutura de proteína em resolução subatômica é auxiliar no entendimento do seu papel biológico, entender as interações com outras proteínas e com ligantes. Os detalhes dessas interações podem ser analisados com maior precisão e, deste modo, podem contribuir para o desenvolvimento de racional de fármacos para tratamento de muitas doenças.

No experimento de difração de Raios X de monocristais de proteínas são medidas as intensidades das ondas espalhadas e não as fases das ondas, o que não permite calcular o mapa de densidade eletrônica de modo direto e construir o modelo tridimensional. Para resolver essa questão há três métodos: Métodos diretos, Substituição Molecular e métodos baseados em sinais anômalos e/ou isomorfos (descrito com maiores detalhes na sessão 5.2.2). Para determinar as

estruturas das proteínas PTP1B e TIPRL $\Delta$ N foram utilizados dois desses métodos. Para a primeira, cujo objetivo era o estudos de interações com ligantes, foi adotado o método mais usual que é a Substituição Molecular, já que existem centenas de estruturas dessa proteína depositadas no banco de dados PDB. Para a segunda, que é uma proteína inédita e se trata de um novo enovelamento, foi adotado o método de sinal anômalo, o que não é usual no país. A estrutura da PTP1B permite estudos no desenvolvimento de fármacos para câncer, diabetes e obesidade. Já a estrutura da TIPRL, resolvida neste trabalho, pode contribuir para o entendimento de interações proteína-proteína e, também, contribuir para o tratamento de doenças.

## 1.7 Objetos

Esta tese teve como objetivo inicial o estudo estrutural das fosfatases PP2A e PTP1B em complexo com inibidores. Os estudos com PTP1B consistiram principalmente em estudos estruturais e ensaios de inibição enzimática utilizando compostos provenientes dos grupos de síntese orgânica do Instituto de Química da Unicamp. Em relação à fosfatase PP2A, os estudos se encaminharam para a resolução da estrutura de sua proteína reguladora TIPRL, que pode ser considerada um inibidor celular proteico.

Mais especificamente, pretendia-se:

- (i) Realizar ensaios de inibição enzimática da PTP1B para verificar a inibição de todos os compostos provenientes dos grupos de síntese orgânica do Instituto de Química da Unicamp;
- (ii) Determinar a estrutura cristalográfica de complexos entre PTP1B e os compostos ativos, gerando informações para estudos adicionais;
- (iii) Expressar e purificar a subunidade catalítica da fosfatase PP2A (PP2Ac) em complexo com sua subunidade reguladora  $\alpha$ 4 e com  $\alpha$ 4 + TIPRL e realizar testes de cristalização com esses complexos e com TIPRL isoladamente;

(iv) Resolver a estrutura cristalográfica de TIPRL e desvendar as bases moleculares da interação de TIPRL com PP2Ac utilizando mutagênese sítio-dirigida, ensaios de interação in vitro e espectrometria de massas.

# **CAPÍTULO 2**

## **Metodología**

## 2. Metodologia

### 2.1 Clones

A síntese do clone da fosfatase PTP1B humana foi realizada, sob encomenda, pela empresa *GenScript* ([www.genscript.com](http://www.genscript.com)) com otimização de códons para expressão de proteína recombinante em bactéria *Escherichia coli* em plasmídeo pUC57. O inserto foi introduzido, por técnicas de subclonagem, em vetor de expressão pET28a. Os clones das outras proteínas alvo dos estudos desta tese foram construídos pela Dra. Juliana Smetana. A Tabela 2.1 traz as informações dos cDNAs clonados que foram usados neste trabalho e a Figura 2.1 o esquema da porção expressa pelas principais proteínas estudadas.

A construção da PTP1B humana corresponde à porção catalítica, solúvel e estável, composta por 298 resíduos de aminoácidos que compõem a porção N-terminal (Tabela 2.2). A deleção dos 137 resíduos de aminoácidos na extremidade C-terminal é referente à porção da proteína que se ancora na membrana do retículo endoplasmático nas células e é composta por resíduos de aminoácidos hidrofóbicos e, estruturalmente, desenovelada. Nesta construção, é adotada uma cauda de histidinas no N-terminal e sítio de clivagem para a proteína trombina bovina (GLVPR-GS).

Foram adotadas duas construções para a proteína TIPRL humana, a primeira TIPRL *Full* corresponde à proteína inteira sem deleções e com cauda de seis histidinas e sítio de clivagem para a protease TEV, *Tabaco Etch Virus*, (ENLYF-Q) em seu N-terminal. A segunda é a TIPRL $\Delta$ N, que possui uma deleção nos 12 primeiros resíduos de aminoácidos na porção N-terminal e, assim como a *Full*, também tem cauda de histidinas e sítio de clivagem de TEV nessa região (Tabela 2.2).

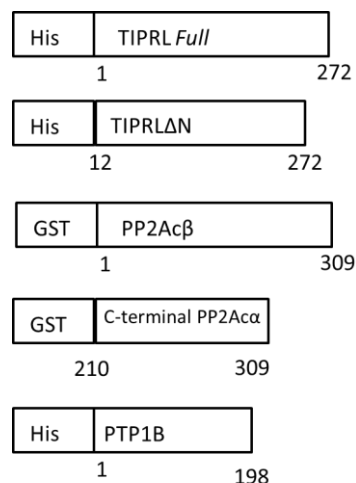
Foram adotadas três construções de PP2Ac nesta tese. A primeira é a PP2Ac $\beta$ , subunidade C isoforma  $\beta$ , fusionada à proteína GST para estudos de *pulldown*. A segunda é a porção C-terminal da PP2Ac, a qual é composta pelos resíduos 210 a 309 e que, também, está fusionada com a GST na região

correspondente ao N-terminal (Smetana e Zanchin 2007). A terceira construção da PP2Ac, é a PP2Ac $\alpha$ 2, isoforma de *splicing* alternativo da isoforma  $\alpha$  sem o éxon número 5 (Migueleti *et. al.*, 2012), a qual forma complexos com as proteínas  $\alpha$ 4 e TIPRL. Esta construção tem cauda de histidinas e sítio de clivagem para a protease TEV em seu N-terminal.

A construção da  $\alpha$ 4 é composta pela porção estruturada no N-terminal do resíduo 1 ao 222; por isso aqui é denominada como  $\alpha$ 4 $\Delta$ 222 (Smetana *et. al.*, 2006). Não possui *tag* de afinidade.

**Tabela 2.1:** Códigos de entrada do GeneBank e UniProt das sequências de nucleotídeos e aminoácidos das fosfatases PTP1B e PP2Ac e as proteínas TIPRL e  $\alpha$ 4 e os vetores onde os cDNA foram clonados.

Proteína	Gene/ GenelD UniProt	Vetor
TIPRL <i>Full</i> / TIPRL $\Delta$ N	TIPRL/261726 O75663	pET-TEV
PP2Ac $\beta$ / C-terminal PP2Ac $\alpha$ / PP2Ac $\alpha$ 2	PPP2CA / 5515 P67775	pET-GEX/ pET-GEX/ pPROEX
$\alpha$ 4 $\Delta$ 222	IGBP1 / 3476 P78318	pET-29a
PTP1B	PTPN1(PTP1B)/ 5770 P18031	pET28-a



**Figura 2.1:** Esquema das construções das principais proteínas estudadas neste trabalho.

**Tabela 2.2:** Sequências de aminoácidos das proteínas TIPRLΔN e PTP1B. As sequências de resíduos de aminoácidos das proteínas usadas nos estudos cristalográficos. A sequência da TIPRLΔN corresponde aos resíduos 12 ao 272 e da PTP1B os resíduos 1 ao 298. Nas porções N-terminal de ambas está localizado a cauda de polihistidina representada em verde e o sítio de clivagem da protease (em vermelho) TEV em TIPRLΔN e trombina em PTP1B.

Proteína	Sequência de Aminoácidos
TIPRLΔN	<p><b>MGHHHHHHENLYF</b>QGHMAS</p> <p><sup>12</sup>FCFGPWKLTASKTHIMKSADVEKLADELHMPSLPEMMFGDNVLRIQHGS            GFGIEFNATDALRCVNNYQGMLKVACAEWQESRTEGEHSKEVIKPYDWT            YTTDYKGTLLGESLKLKVVPTTDHIDTEKLEKAREQIKFFEEVLLFEDELH            DHGVSSLSVKIRVMPSSFALLRFFLRIDGVLIRMNDTRLYHEADKTYML            REYTSRESKISSLMHVPPSLFTEPNEISQYLPIKEAVCEKLIFFERIDPN            PADSQKSTQVE<sup>272</sup></p>
PTP1B	<p><b>MGSSHHHHHHSSGLVPR</b>GSHMEFM</p> <p><sup>1</sup>MEMEKEFEQIDKSGSWAAIYQDIRHEASDFPCRVAKLPKNKNNRNRDVS            PFDHSRIKLNQEDNDYINASLIKMEEAQRSYILTQGPLPNTCGHFWEMVW            EQKSRGVVMLNRVMEKGLKCAQYWPQKEEKEMIFEDTNLKLTLISEDIK            SYTTVRQLELENLTTQETREILHFHYTTWPDFGVPEPASPFLNLFKQVRE            SGLSLSPEHGPPVVVHCSAGIGRSGTFCLADTCLLLMDKRKDPSSVDIKKVL            LEMRKFRMGLIQTDQLRFSYLAVIEGAKFIMGDSSVQDQWKELSHED<sup>298</sup></p>

## 2.2 Expressão de proteínas em sistema recombinante

### 2.2.1 Expressão em meio LB líquido

Células competentes de *Escherichia coli* foram submetidas ao procedimento de transformação com o plasmídeo adequado, usando-se protocolo de transformação por choque térmico. Placas de cultura foram inoculadas com as bactérias transformadas, usando meio LB sólido com antibiótico para a seleção de colônias transformadas.

Uma colônia transformada e selecionada de acordo com resistência a antibiótico foi transferida para um tubo de pré-inóculo, contendo meio de cultura LB líquido e antibiótico. A cultura foi mantida *overnight* a 37°C sob agitação constante. Em seguida, o pré-inóculo foi transferido para um *erlenmeyer* contendo meio LB líquido e antibiótico. A cultura foi crescida a 37 °C por 3 horas, até atingir um número de células suficientes para a indução da expressão da proteína de interesse (ou seja, quando a DO, medida pela absorbância a 600 nm estiver entre 0,5 e 0,8 U.A.). Atingida a DO, foi adicionado IPTG ou lactose ao meio para induzir a expressão do gene que codifica as proteínas de interesse. A cultura foi mantida sob agitação constante de 200 rpm, a temperatura e o tempo foram controlados, conforme dados



na Tabela 2.3. Para chegar a esses protocolos de expressão, foram realizados ensaios nos quais cada parâmetro foi testado e analisado.

**Tabela 2.3:** Condições de expressão das proteínas recombinantes.

Proteína Ou complexos	Cepa de <i>E. coli</i>	Antibiótico	Indução	Temperatura (°C)	Tempo de expressão (horas)
TIPRL	BL21 (DE3)	Canamicina	Lactose 14 mmol.L <sup>-1</sup>	25	5
GST- PP2A $\alpha$ 2: TIPRL <i>Full</i>	BL21 (DE3)	Canamicina Ampicilina	IPTG 0,5 mmol.L <sup>-1</sup>	25	5
PP2A $\alpha$ 2: $\alpha$ 4 $\Delta$ 222	BL21 pRARE	Canamicina Ampicilina	IPTG 0,1 mmol.L <sup>-1</sup>	30	20
TEV	BL21 (DE3)	Canamicina	IPTG 0,5 mmol.L <sup>-1</sup>	37	5
PTP1B	BL21 (DE3)	Canamicina	IPTG 0,5 mmol.L <sup>-1</sup>	25	5

### 2.2.2 Expressão de TIPRL $\Delta$ N marcada com Selenometionina

A expressão em Meio Mínimo é feita em um meio de cultura controlado, composto por sais inorgânicos (meio M9), glicose e suplementos com tiamina e aminoácidos essenciais como lisina, valina, isoleucina, fenilalanina, dentre outros, e com metionina marcada com selênio – selenometionina ou Se-Met. Desse modo, a bactéria sintetiza somente proteína com a selenometionina sem o risco de incorporar aminoácidos não marcados. Realizou-se a expressão com IPTG 0,5 mmol.L<sup>-1</sup> em meio mínimo com a Se-Met como descrito na literatura (Studier 2005). A proteína expressa com Se-Met foi digerida por tripsina (16 horas a 37 °C) e analisada por espectrometria de massas para confirmar a incorporação da selenometionina.

### 2.3 Purificação das proteínas

As proteínas e complexos de interesse (TIPRL, PTP1B, PP2A $\alpha$ 2: $\alpha$ 4 e TEV) foram produzidas com uma cauda de histidina em seu N-terminal. As células de bactéria foram separadas do meio de cultura por centrifugação (3500 rpm, 15

minutos a 4°C). Estas células foram lisadas através de lise enzimática, utilizando a enzima lisozima (0,1 mg.mL<sup>-1</sup>) e ultrassom (10 a 20 ciclos de 15 segundos). Os restos celulares foram separados por nova centrifugação (20.000 rpm, 40 minutos a 4 °C) e o sobrenadante, contendo a proteína de interesse, foi submetido às etapas de purificação.

O sobrenadante proteico foi incubado por 1 hora com a resina de níquel (Ni Sepharose™ 6 Fast Flow – GE Healthcare) para purificação IMAC – cromatografia de afinidade por metal imobilizado. A resina foi equilibrada e lavada para a retirada de proteínas com baixa afinidade ao níquel. Em seguida, foi realizada a eluição das proteínas com afinidade ao níquel, por lavagem da resina com solução contendo concentrações crescentes de imidazol. Os tampões utilizados para cada proteína podem ser vistos na Tabela 2.4.

Para formar o complexo PP2A $\alpha$ 2: $\alpha$ 4 $\Delta$ 222:TIPRL $\Delta$ N, as expressões do complexo PP2A $\alpha$ 2: $\alpha$ 4 $\Delta$ 222 e da proteína TIPRL $\Delta$ N foram independentes. Após a lise celular, os extratos dessas duas amostras foram misturados na proporção de 5:1, incubado em resina de cobalto (Talon® Superflow™ - GE Healthcare) prosseguindo para a etapa de purificação. A purificação foi feita equilibrando e lavando a resina e, posteriormente, eluindo-a com imidazol.

Após a purificação em IMAC, a proteína ou complexo de interesse teve sua cauda de histidina clivada por uma protease específica (TEV na proporção de 0,1 mg por mg de proteína ou complexo ou trombina na proporção de 0,5U por mg de proteína; ambas as clivagem foram *overnight* a 4 °C) e uma nova etapa de purificação, por exclusão molecular, foi utilizada. A cromatografia de exclusão molecular foi realizada em coluna contendo resina Superdex75 (GE Life Science) acoplada a sistema de cromatografia líquida de alta eficiência ÄKTA (ÄKTA Purifier – GE Life Sciences).

**Tabela 2.4:** Relação dos tampões utilizados para os processos de purificação de cada proteína.

Proteínas	Tampão Equilíbrio	Lavagem	Gradiente de	
			imidazol mmol.L <sup>-1</sup>	Tampão Gel Filtração
TIPRL	Tris, pH 7,5, 100 mmol.L <sup>-1</sup> NaCl 300 mmol.L <sup>-1</sup> $\beta$ -mercaptoetanol 10mmol.L <sup>-1</sup> e Imidazol 20 mmol.L <sup>-1</sup>	Tris pH 7,5, 100 mmol.L <sup>-1</sup> NaCl 1 mol.L <sup>-1</sup> e $\beta$ -mercaptoetanol 10 mmol.L <sup>-1</sup>	400 e 500	Tris pH 7,5 10 mmol.L <sup>-1</sup> NaCl 150 mmol.L <sup>-1</sup> e $\beta$ -mercaptoetanol 10 mmol.L <sup>-1</sup>
Complexo PP2A $\alpha$ 2: $\alpha$ 4 $\Delta$ 222 e PP2A $\alpha$ 2: $\alpha$ 4 $\Delta$ 222: TIPRL $\Delta$ N	PBS pH 7,4 Glicerol 5% e Imidazol 20 mmol.L <sup>-1</sup>	PBS pH 7,4 e Glicerol 5%	200	Tris pH 7,5 10 mmol.L <sup>-1</sup> NaCl 100 mmol.L <sup>-1</sup> e DTT 0,5 mmol.L <sup>-1</sup>
PTP1B	Tris pH7,5, 100 mmol.L <sup>-1</sup> NaCl 300 mmol.L <sup>-1</sup> DTT 5mmol.L <sup>-1</sup> e Imidazol 30 mmol.L <sup>-1</sup>	Tris pH 7,5, 100 mmol.L <sup>-1</sup> NaCl 500 mmol.L <sup>-1</sup> e DTT 5mmol L <sup>-1</sup>	100, 250 e 500	Tris pH 7,5 10 mmol.L <sup>-1</sup> NaCl 150 mmol.L <sup>-1</sup> DTT 1 mmol.L <sup>-1</sup> e EDTA 0,3 mmol.L <sup>-1</sup>
TEV	Tris pH 8,0, 50 mmol.L <sup>-1</sup> NaCl 200 mmol L <sup>-1</sup> Glicerol 10% e Imidazol 20 mmol.L <sup>-1</sup>	Tris pH 8,0 50 mmol.L <sup>-1</sup> e NaCl 1 mol.L <sup>-1</sup>	300 e 500	Esta etapa da purificação não foi necessária

A quantificação da proteína purificada foi feita por medidas de absorbância a 280 nm, usando a Lei de Lambert-Beer e os coeficientes de extinção molar  $\epsilon$ , das proteínas purificadas calculados pelo servidor *ExpASy Proteomics Server* ([www.expasy.org](http://www.expasy.org)).

As etapas de expressão e purificação das proteínas de interesse foram acompanhadas por análises de migração da proteína em gel de 12 % (v/v) poliacrilamida em condições desnaturantes (SDS/PAGE) (Laemmli 1970). Por SDS/PAGE pode-se checar a massa aproximada de espécies monoméricas da proteína purificada (Shapiro *et. al.*, 1967) através da análise concomitante de

proteínas com massas moleculares conhecidas, além de ser possível, também, checar a pureza da amostra.

## 2.4 Análises de proteínas por CD e DLS

Após a purificação, as proteínas foram analisadas por espectroscopia de dicroísmo circular (CD) e por espalhamento dinâmico de luz (DLS), com a finalidade de confirmar sua integridade estrutural e conteúdo de estrutura secundária, bem como a monodispersidade da amostra. Ambas as técnicas foram importantes ferramentas no controle de qualidade da proteína produzida, precedendo os estudos subsequentes. Os dados foram coletados em equipamentos disponíveis no IQ/UNICAMP e no LEC (LNBio/CNPEM – Campinas SP).

Os espectros de CD na região do UV distante (190-260 nm) foram medidos num espectropolarímetro JASCO J – 715 (JASCO) equipado com o módulo Peltier para o controle de temperatura disponível no IQ/UNICAMP, utilizando uma cubeta de quartzo de 1 mm de caminho óptico. A concentração da proteína TIPRLΔN e complexos PP2Aα2:α4Δ222 e PP2Aα2:α4Δ222:TIPRLΔN foi de 0,1 mg.mL<sup>-1</sup> dissolvido em tampão de fosfato de sódio 20 mmol.L<sup>-1</sup>, pH 7,3, cada espectro obtido foi o resultado de 16 acumulações medidas a 20 °C. Os resultados foram expressos como elipticidade molar ( $\theta$  = graus.cm<sup>2</sup>.dmol<sup>-1</sup>).

A coleta de dados da proteína fosfatase PTP1B foi realizada em concentração de 0,1 mg.mL<sup>-1</sup> e em diferentes tampões na concentração de 20 mmol.L<sup>-1</sup>: glicina-HCl pH 2,5 e 3,0; acetado de sódio pH 4,0; 4,5; 5,0 e 5,5; MÊS pH 6,0; Bis-Tris propano pH 6,5; 7,0; 7,5; 8,0 e 9,0. O objetivo deste experimento foi encontrar o pH no qual a proteína estivesse com mais estrutura secundária e assim empregar esse valor de pH nos ensaios de cinética. Além disso, foi realizado o experimento de desnaturação térmica com a proteína fosfatase PTP1B. A amostra foi aquecida de 20 a 90 °C e resfriada até 20 °C com uma taxa de aquecimento ou resfriamento de 1 °C.min<sup>-1</sup>, e monitoradas no UV distante. Os espectros foram registados a 222 nm, valor correspondente ao sinal advindo de α-hélices na estrutura secundária. Essa medida foi realizada para encontrar a faixa de

temperatura na qual a estrutura fosse mantida mediante ao aquecimento e, assim, guiar os ensaios de cinética.

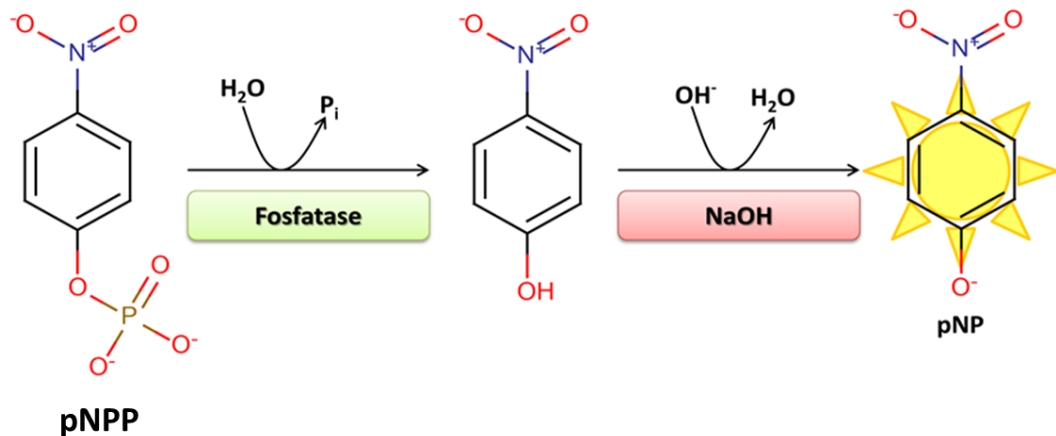
Para verificar os parâmetros hidrodinâmicos e homogeneidade de proteínas em solução, as amostras foram analisadas por espalhamento dinâmico de luz (DLS). Os experimentos foram realizadas no Laboratório de Espectroscopia e Calorimetria, LEC (LNBio) no equipamento Dynapro 810 (Protein Solutions). A concentração das amostras de TIPRLΔN e complexos PP2Aα2:α4Δ222 e PP2Aα2:α4Δ222:TIPRLΔN foi de 2,5 mmol.L<sup>-1</sup>. Antes de cada medida as amostras foram centrifugadas a 14.000 rpm, 15 minutos a 4 °C e, posteriormente, inseridas na cubeta de 80 μL. O equipamento é equipado com o módulo Peltier para o controle de temperatura, e as leituras foram realizadas a 18 °C. O comprimento de onda e potência do laser foram de 830 nm e 80%, respectivamente. A quantidade de aquisições realizadas foi de 100, com o tempo de 5 s. Os parâmetros hidrodinâmicos foram determinados usando o *software* DYNAMICS v.6.10.1.2.16. O objetivo desse estudo foi analisar a polidispersidade das amostras antes de encaminhá-las à triagem de condições de cristalização disponíveis no ROBOLAB. Este ensaio não foi realizado para a PTP1B.

## **2.5 Ensaios enzimáticos e Triagem de compostos para a proteína PTP1B**

Os compostos sintetizados pelos grupos de Química Orgânica do Instituto de Química/Unicamp participantes no projeto Temático: "Biologia Química: novos alvos moleculares naturais e sintéticos contra o câncer. Estudos estruturais, avaliação biológica e modo de ação" (processo FAPESP: 09/51602-5) foram submetidos a triagens conforme o seguinte protocolo. As leituras de dados foram feitas em espectrofotômetro SpectraMax M2e (Molecular Devices).

Os ensaios foram realizados em larga escala, em microplacas de 96 poços, para triagem das pequenas moléculas com atividade inibitória da PTP1B, através de medidas da atividade da enzima na presença ou na ausência dos inibidores. A mistura de reação (95 μL) contendo tampão de acetato de sódio pH 5,5, 100 mmol L<sup>-1</sup> a enzima (9,1 nmol.L<sup>-1</sup>) e o composto de interesse (5 mmol.L<sup>-1</sup>) foram incubados durante 10 min a 37 °C. A reação foi iniciada pela adição de 5 μL do

substrato sintético pNPP (para-nitrofenil-fosfato), em concentração final de  $1 \text{ mmol.L}^{-1}$  (Figura 2.2). Após 10 minutos de reação, esta foi interrompida pela adição de  $100 \mu\text{L}$  de  $\text{NaOH } 1 \text{ mol.L}^{-1}$ . A absorbância relativa foi lida a  $405 \text{ nm}$  uma vez que o pNP (produto da reação da proteína com o pNPP) desprotona-se em pH elevado (efeito causado pelo hidróxido de sódio), ficando amarelo .



**Figura 2.2:** Esquema do mecanismo de reação da PTP1B com o substrato sintético pNPP.

## 2.6 Cristalografia de Proteínas

### 2.6.1 Cristalização

Os ensaios de cristalização da TIPRLAN foram conduzidos utilizando o método de gota sentada. A concentração de proteína utilizada variou de  $70$  a  $100 \text{ mg.mL}^{-1}$  em tampão Tris  $20 \text{ mmol.L}^{-1}$  pH  $7,5$ ; cloreto de sódio  $25 \text{ mmol.L}^{-1}$  e  $\beta$ -mercaptoetanol  $10 \text{ mmol.L}^{-1}$ . Os cristais crescem em duas condições, descritas abaixo, em aproximadamente de 10 dias. Essas condições foram obtidas através de ensaios de cristalização em matriz esparsa (Cudney *et. al.*, 1994) no ROBOLAB (LNBio/CNPEM, Campinas - SP). Após refinamento, as condições otimizadas são as seguintes:

1)  $18 \text{ }^\circ\text{C}$  em solução precipitante que contém Fosfato de sódio di-potássio  $0,1 \text{ mol.L}^{-1}$  pH  $6$  a  $6,5$ ; Cloreto de sódio  $1,5$  a  $3 \text{ mol.L}^{-1}$ .

2) 18 °C em solução precipitante que contém tampão SPG 0,1 mol.L<sup>-1</sup> pH 5,5 a 6,7; Cloreto de sódio 1,5 a 3 mol.L<sup>-1</sup>.

A cristalização da PTP1B foi conduzida através do método de gota sentada. A concentração da proteína de PTP1B utilizada foi de 7 a 10 mg.mL<sup>-1</sup> em tampão Tris 10 mmol.L<sup>-1</sup> pH 7,5; cloreto de sódio 25 mmol.L<sup>-1</sup> e DTT 3 mmol.L<sup>-1</sup> (Barford *et. al.*, 1994). Os cristais crescem a 4 °C em solução precipitante que contém tampão Hepes 0,1 mol.L<sup>-1</sup> pH 7,5, acetato de magnésio 0,2 mol.L<sup>-1</sup>, 10-20% PEG 8000, conforme descrito na literatura (Buhrman *et. al.*, 2005; Fauman *et. al.*, 1998; Reynolds *et. al.*, 1999; Zabell *et. al.*, 2006; Zhang *et. al.*, 1998). A concentração de inibidor usada foi de 15 vezes a concentração de proteína no caso de cocristalização com o composto de interesse. Além disso, após crescimento do cristal apo (sem ligante), foi feito *soaking* (banho) na solução precipitante enriquecida com 50% (v/v) dos ligantes.

## 2.6.2 Coleta de dados e obtenção da estrutura tridimensional

Após a obtenção de cristais das proteínas, testes iniciais de difração foram conduzidos no laboratório institucional do IQ/UNICAMP, em difratômetro Bruker Kappa APEX II DUO (fonte de Cu e Mo), utilizando-se fonte de microfoco com anodo de cobre (Quazer MX) e detector CCD Figura 2.3. Como solução crioprotetora, utilizou-se 20% (v/v) 2,3-butanediol ou MDP para os cristais de TIPRLAN ou 16% (v/v) 2,3-butanediol para PTP1B. Os conjuntos de dados de difração foram coletados nesse equipamento e na linha de difração W01B- MX2 (Guimaraes *et. al.*, 2009) do LNLS/CNPEM, Campinas-SP, equipada com detector CCD, ambas as coletas com comprimento de onda de  $\lambda = 1,5417\text{\AA}$ . Os cristais de TIPRLAN marcados com Se-Met foram coletados de modo remoto na linha I02 – Diamond Light Source/Inglaterra, detector bidimensional Pilatus 6M e  $\lambda = 0,99\text{\AA}$  (borda de absorção do selênio). Os dados foram coletados pelo método de rotação (Z. Otwinowski 1997) e o cristal foi mantido a 100 K durante toda a coleta para minimizar os danos por exposição à radiação. Os parâmetros de coleta foram ajustados a fim de obter-se a maior resolução e qualidade dos dados possível.

O padrão de difração foi indexado e as intensidades das reflexões hkl nas imagens de difração foram integradas e escalonadas usando os programas *Mosflm/Scala* (Leslie 1999, 2006) para os conjuntos coletados na linha de luz do LNLS; pacote de programas Bruker para as coletas *no* IQ/Unicamp. Já na coleta de dados realizada no Diamond Light Source/Inglaterra, a estratégia de coleta foi calculada com os programas EDNA (Incardona *et. al.*, 2009) e *Mosflm*; a linha conta com um processamento automatizado usando os programas *Xia2* (Winter *et. al.*, 2013) e *FAST\_DP* (Winter e McAuley 2011). O faseamento automatizado é feito pelo *FAST\_EP* (Winter e McAuley 2011). A localização dos átomos de selênio foi feita pelo conjunto de programas *ShelxC/D/E* (Sheldrick 2010) e o modelo foi constituído pelo programa *ARP/wARP* (Langer *et. al.*, 2013).

A estrutura cristalográfica da PTP1B encontrava-se disponível no PDB (códigos: 3A5J e 1AAX), e estas coordenadas atômicas das estruturas foram utilizadas como ponto de partida para obtenção das fases iniciais, usando o método de substituição molecular com o programa *Phaser* (McCoy *et. al.*, 2007), pacote *CCP4* (The CCP4 suite: programs for protein crystallography 1994). A construção e refinamento dos modelos atômicos foram realizados com o programa *Coot* (Emsley e Cowtan 2004) e o *software* de refinamento *Refmac5* (Murshudov *et. al.*, 1997), implementado no pacote *CCP4* (The CCP4 suite: programs for protein crystallography 1994).

No caso da TIPRLAN, não havia estruturas de homólogas disponíveis. Ainda assim, esforços exaustivos foram feitos, sem sucesso, utilizando-se o estado da arte em modelagem molecular (I-TASSER (Zhang 2008), Robetta (Rohl *et. al.*, 2004), entre outros) e programas para a Substituição Molecular (BALBES (Long *et. al.*, 2008), Mr. BUMP (Keegan e Winn 2008), entre outros). A estrutura foi resolvida utilizando-se a técnica de SAD (*Single-wavelength Anomalous Dispersion*), a partir do espalhamento anômalo dos cristais de proteína marcada com Se-Met.





**Figura 2.3:** Difratorômetro Bruker Kappa APEX II Duo do Instituto de Química/Unicamp.

## 2.7 SAXS

### 2.7.1 Coleta de dados

Os experimentos de espalhamento de Raios X a baixo ângulo (SAXS) foram realizados na linha de luz D01A-SAXS1 do LNLS (Laboratório Nacional de Luz Sincrotron) com detector bidimensional Pilatus. As amostras da proteína TIPRLΔN ( $0,5 - 25 \text{ mg mL}^{-1}$ ) e seus tampões foram centrifugados por 10 minutos, a 10.000 rpm e em temperatura ambiente para eliminar possíveis agregados. Uma amostra de, aproximadamente,  $120 \text{ } \mu\text{l}$  foi cuidadosamente injetada no porta amostras constituído por duas janelas paralelas de mica ou em capilar de vidro. Os dados foram coletados utilizando comprimento de onda de  $1,55 \text{ \AA}$  e distância entre a amostra e o detector de  $1603,43 \text{ mm}$ , levando a um intervalo  $0,0067 \leq q \leq 0,278 \text{ \AA}^{-1}$  onde  $q$  é o módulo do vetor de espalhamento, definido como  $q = 4\pi\text{sen}(\theta)/\lambda$ , no qual  $\theta$  é metade do ângulo de espalhamento.

Amostras de proteína e da solução tampão (usado como branco na análise dos dados) foram expostas aos Raios X com 5 *frames* de 5 segundos com intervalos de 2 segundos entre cada curva inicialmente e os outros *frames* variam conforme a concentração da amostra e os dados a radiação. As curvas foram integradas radialmente e normalizadas pelo programa FIT2D (Moy *et. al.*, 1996).

Os tampões das amostras medidas no SAXS foram: (i) Fosfato de sódio  $20 \text{ mmol.L}^{-1}$  pH7,4; NaCl  $50 \text{ mmol.L}^{-1}$   $\beta$ -mercaptoetanol  $5 \text{ mmol.L}^{-1}$ ;

(ii) Tris 20 mmol.L<sup>-1</sup> pH7,5; NaCl 25 ou 50 mmol.L<sup>-1</sup> e β-mercaptoetanol variando de 10 a 50 mmol.L<sup>-1</sup>.

### 2.7.2 Análise das curvas

A análise de Guinier foi realizada para obter o raio de giro da macromolécula ( $R_g$ ), a intensidade o espalhamento no ângulo zero  $I(0)$ , assumindo a lei de Guinier, onde:  $I(q) = I(0)\exp(-q^2[(R_g^2)/3])$  para ângulo perto de zero ( $q < 1,3/R_g$ ). Os cálculos mais precisos para tais valores foram realizados pelo programa GNOM (Svergun 1992), através da transformada inversa de Fourier. Com isso foram obtidas as informações de função de distribuição de distâncias  $p(r)$ , dimensão máxima da molécula ( $D_{max}$ ) e sua anisometria. O gráfico de Kratky ( $Iq^2 \times q$ ) foi plotado para obter informações a respeito do correto enovelamento da proteína (Putnam *et. al.*, 2007).

## 2.8 Cross-linking acoplado a MS

Os experimentos de espectrometria de massas foram realizados em colaboração com o Prof. Dr. Fábio César Gozzo do Instituto de Química/Unicamp, e sua orientada de doutorado Ma. Tatiani B. Lima.

### 2.8.1 Reação de ligação cruzada utilizando DSS

As proteínas TIPRL *Full* e TIPRL $\Delta$ N purificadas em tampão fosfato de sódio pH 7,5, 50 mmol.L<sup>-1</sup>, cloreto de sódio 200 mmol.L<sup>-1</sup>, glicerol 5% e β-mercaptoetanol 10 mmol.L<sup>-1</sup> com concentração estimada de 15 mg.mL<sup>-1</sup> foram submetidas à reação de ligação cruzada. O agente de ligação cruzada DSS (suberato de N-hidroxisuccinimidila) foi dissolvido em DMF (N,N-dimetilformamida) anidro de modo a obter uma concentração de 10 mg.mL<sup>-1</sup> ou 27,1 mmol.L<sup>-1</sup>. Uma alíquota desta solução foi adicionada à solução de proteína de modo a obter um excesso molar de 100 vezes de DSS, mantendo a concentração de DMF menor que 10% (v/v). A reação de ligação cruzada foi realizada por duas horas a temperatura

ambiente sob leve agitação. A reação foi cessada adicionando 100  $\mu\text{L}$  de bicarbonato de amônio 50  $\text{mmol.L}^{-1}$ . A amostra foi submetida à redução de pontes de dissulfeto por adição de 15  $\mu\text{L}$  de solução de ditioneitol (DTT) 10 mM em tampão bicarbonato de amônio 50  $\text{mmol.L}^{-1}$ , incubando-se por 30 minutos a 60 °C. A subsequente alquilação dos grupos tiol livres foi realizada adicionando-se 10  $\mu\text{L}$  de solução de iodoacetamida 50  $\text{mmol.L}^{-1}$  em tampão bicarbonato de amônio 50  $\text{mmol.L}^{-1}$ , por 30 minutos em temperatura ambiente. Para digestão enzimática, foi utilizada tripsina porcina (Sequencing Grade Modified Trypsin, Promega) em uma proporção de 1:50 (m:m) enzima:substrato, incubando-se por 16 horas a 37 °C.

Após digestão enzimática, a mistura de peptídeos resultantes foi fracionada utilizando cartuchos de extração em fase sólida (SPE) Oasis HLB (Waters). A amostra foi eluída em cinco frações utilizando como eluentes misturas de acetonitrila/água nas proporções 10:90, 15:85, 20:80, 30:70 e 80:20, respectivamente.

As amostras foram analisadas por LC-MS/MS, utilizando um cromatógrafo a líquido *nanoAcquity* UPLC (Waters), equipado com coluna *Acquity* BEH C18 (100  $\mu\text{m}$  x 100 mm, partículas de 1,7  $\mu\text{m}$ ), acoplado a um espectrômetro de massas Waters Synapt HDMS (Waters). Foi injetado 4,5  $\mu\text{L}$  de cada fração (40 pmol de peptídeos por injeção).

As corridas de LC-MS/MS foram processadas pelo *software* MASCOT Distiller v. 2.4 (Matrix Science Ltd) e analisadas por busca em banco de dados empregando-se o sistema MASCOT v.2.4. (Matrix Science Ltd). Para avaliação de presença de ligações cruzadas intra- e inter-peptídeos, os dados pré-processados no MASCOT Distiller foram submetidos aos *softwares* *Crux* e *SIM-XL*.

## 2.9 Troca de Hidrogênio-Deutério HDX Análise por espectrometria de Massas

Para análise por troca hidrogênio-deutério foram preparadas: (i) solução estoque de TIPRL *Full* de concentração 99,4  $\mu\text{mol.L}^{-1}$  em tampão Tris.HCl, pH 7,5, 20  $\text{mmol.L}^{-1}$ , NaCl 100  $\text{mmol.L}^{-1}$ ,  $\beta$ -mercaptoetanol 10  $\text{mmol.L}^{-1}$  (ii) tampão

deuterante (Tris.DCl, pD 7,5, 20 mmol.L<sup>-1</sup> NaCl, 100 mmol.L<sup>-1</sup>, β-mercaptoetanol 10 mmol.L<sup>-1</sup>); (iii) tampão *quench* em água milli-Q (cloridrato de guanidina 800 mmol.L<sup>-1</sup>, ácido fórmico 0,8%, DTT 20 mmol.L<sup>-1</sup>, pH 2,5); (iv) solução estoque do peptídeo DYFL de concentração 35,95 mmol.L<sup>-1</sup> dissolvido em dimetilformamida (DMF); (v) solução estoque do peptídeo DpYFL de concentração 31,44 mmol.L<sup>-1</sup> dissolvido em DMF. Os peptídeos DYFL e DpYFL, que são o extremo C-terminal da proteína PP2Ac, foram sintetizados pela empresa Proteimax com grau de pureza maior que 95% e rendimento de 25,5 e 27,7 mg respectivamente. Para as reações de troca hidrogênio-deutério a proteína TIPRL *Full* foi incubada com os peptídeos por 2 horas com proporção (1:10) (proteína:peptídeo) e o padrão TIPRL *Full* foi incubado com o mesmo volume de DMF utilizado nos ensaios com os peptídeos.

As reações de troca hidrogênio-deutério foram realizadas diluindo 15 vezes as diferentes soluções de TIPRL *Full* (TIPRL *Full* padrão, TIPRL *Full* com peptídeo DYFL e TIPRL *Full* com peptídeo DpYFL) em tampão deuterante em temperatura ambiente. Após os tempos de reação, as reações de troca foram paradas com tampão *quench*, a 4 °C, com volume igual ao da reação de troca, e em seguida, as amostras foram injetadas. A cinética de troca hidrogênio-deutério foi realizada em diferentes tempos de troca (controle – 0s, 10s, 1 min, 10 min, 1h e 2h).

As amostras foram injetadas em um sistema *nanoAcquity* UPLC com Tecnologia HDX acoplado ao Synapt G1 HDMS (Waters Corporation, USA). A digestão *online* foi realizada utilizando uma coluna de pepsina imobilizada (2x30 mm, Applied Biosystem, USA) por 3 minutos a 15 °C com o fluxo de 50 μL.min<sup>-1</sup>. Os peptídeos gerados foram dessalinizados *online* em uma pré-coluna (ACQUITY UPLC BEH C18 1,7 μm VanGuard, Waters) a 0°C e separados por uma coluna analítica (ACQUITY UPLC BEH C18 1,7 μm, 1 mm x 100 mm, Waters) a 0 °C com fluxo de 50 μL.min<sup>-1</sup>. As corridas foram processadas pelo *Protein Lynx Global Server* v.2.4. (Waters Co.) e *DynamX* v.3.0 (Waters Co.).

## 2.10 Mutagênese Sítio-Dirigida

As mutações foram selecionadas com base na estrutura do complexo TIPRL-peptídeo e os *primers* foram desenhados utilizando a ferramenta disponível no endereço eletrônico <http://www.genomics.agilent.com/primerDesignProgram.jsp> (Agilent) e sintetizados pela empresa Exxtend com grau de pureza RP-OPC. As reações de mutagênese consistiam em um volume total de 25  $\mu\text{L}$  contendo 150 ng de cada um dos primers descritos na Tabela 2.5 e 2.6, 140 ng de DNA molde, 2  $\mu\text{L}$  de DNTp 5  $\text{mmol.L}^{-1}$ , 1  $\mu\text{L}$  de  $\text{MgSO}_4$  50  $\text{mmol.L}^{-1}$ , 1  $\mu\text{L}$  da enzima *Pfu* polimerase (Thermo Scientific) e 2,5  $\mu\text{L}$  do tampão 10x que acompanha a enzima. As reações foram submetidas a uma desnaturação inicial (95 °C por 30 segundos) seguida de 18 ciclos de desnaturação (95 °C por 30 segundos), anelamento (60 °C por 30 segundos) e extensão (68 °C por 7 minutos) em termociclador. Para verificar a eficiência da reação, 5  $\mu\text{L}$  foram analisados em gel de agarose 0,8 % corado com GelRed. Em seguida, as reações foram tratadas com 10 unidades da enzima *DpnI* (Thermo Scientific) por uma hora a 37 °C. Essa enzima de restrição reconhece e cliva a sequência GATC apenas quando metilada na adenina, eliminando, dessa forma, o DNA parental não mutado. Após o tratamento com *DpnI*, as reações foram transformadas em *E. coli* DH5 $\alpha$  utilizando o método do choque térmico, com 150  $\mu\text{L}$  de células competentes e 20  $\mu\text{L}$  da reação. As colônias foram selecionadas em meio LB com antibiótico apropriado (ampicilina 50  $\mu\text{g.mL}^{-1}$  para as mutações a partir dos vetores pGEX-PP2Ac $\beta$  e pGEX-PP2Ac $\alpha$ C-Terminal, e canamicina 50  $\mu\text{g.mL}^{-1}$  para aquelas provenientes do vetor pET-TEV-TIPRL). As placas foram incubadas na estufa a 37 °C por 16 horas. Para cada reação, cerca de 3 a 5 transformantes foram inoculados em 10 mL de meio LB líquido com antibiótico apropriado. As culturas foram incubadas por 16 horas a 37 °C e agitação de 200 rpm e em seguida utilizadas para extração de DNA plasmidial com o kit de miniprep Qiagen. Os vetores foram analisados por sequenciamento para verificar a correta incorporação da mutação.

**Tabela 2.5:** Primers utilizados para as mutações pontuais na TIPRL *Full*.

Mutantes TIPRL <i>Full</i>	Sequência (5' - 3')
N68A-F	catcttaacgcatctgtagcagcgaactcaattccaaagccaga
N68A-R	tctggcctttggaattgagttcgctgctacagatgcgttaagatg
D71L-F	gttggtttacacatcttaacgctaataatgtagcattgaaactcaattccaaagccaga
D71L-R	tctggcctttggaattgagttcaatgctacattagcgttaagatggtgtaaacaac
I136T-F	ctctggcctttcaatttttctgtatccgtatgatctggttaggtacaacctt
I136T-R	aagggtgtacctacaacagatcatacggatacagaaaaattgaaagccagag
L141A-F	cttaatctggttctctggcctttcgctttttctgtatctatatgatctggttag
L141A-R	ctacaacagatcatatagatacagaaaaagcgaagccagagaaacagattaag
I147A-F	aaaggagaacttcttcaaaaaacttagcctgttctctggcctttcaatttttctg
I147A-R	cagaaaaattgaaagccagagaaacaggttaagtttttgaagaagtctccttt
F150A-F	aaaaggagaacttcttcagcaaaacttaatctggttctctggcctttcaatttttctg
F150A-R	cagaaaaattgaaagccagagaaacagattaagtttgctgaagaagtctcctttt
F180A-F	ccgcaacagcagggcgaagctagaaggcattactctaatac
F180A-R	gattagagtaatgccttctagctttgccctgctggtgctg
L182A-F	ctcaagaaaaaccgcaacgccaggaaaaagctagaaggca
L182A-R	tgcttctagcttttttctggcgttgctggttttttcttgag
R184A-F	ccatcaattctcaagaaaaacgccaacagcaggaaaaagctaga
R184A-R	tctagctttttcctgctggttgccgttttttcttgagaattgatgg
R200A-F	ttgtcagcctcatggtaaaagtgccgtgtcattcattctgataag
R200A-R	cttatcagaatgaatgacacggcactttaccatgaggctgacaa
S239E-F	cttcctttattggtaaatactgctctattttcattaggttccgtgaagaggggaag
S239E-R	cttcctcttcacggaaacctaatgaaatagagcagtatttaccataaaggaag
F180L182AA-F	tctcaagaaaaaccgcaacgccaggcgaagctagaaggcattactctaattctcaca
F180L182AA-R	tgtgaagattagagtaatgccttctagctttgccctggcgttgctggtttttcttgaga

**Tabela 2.6:** Primers utilizados para as mutações pontuais na PP2Ac $\beta$  e C-terminal da PP2Ac.

Mutantes PP2Ac	Sequência (5' - 3')
C-Terminal Y307E-F	aggagaaatattataggaactcgtctgggggtgcccgtg
C-Terminal Y307E-R	cacggcgcaccccagacgagttcctataaatttctcct
C-Terminal T304Stop-F	aattttataggaagtagtctggcctagcgcctgtaacatgaggctc
C-Terminal T304Stop-R	gagcctcatgtttacacggcgttagccagactacttctataaatt
PP2Ac $\beta$ Y307E-F	attggatccttacaggaactcgtctgggggtacgacgag
PP2Ac $\beta$ Y307E-R	ctcgtcgtaccccagacgagttcctgtaaggatccaat
PP2Ac $\beta$ T304Stop-F	cttacaggaagtagtctggcctaacgacgagtaacatgtggctc
PP2Ac $\beta$ T304Stop-R	gagccacatgttactcgtcgttagccagactacttctgtaag

## 2.11 GST-Pulldown

As células bacterianas *E. coli* BL21 (DE3) foram cotransformadas com os vetores que codificam PP2Ac (isoforma PP2Ac $\beta$  ou C-terminal da PP2Ac) e TIPRL *Full*. As células foram cultivadas em meio LB líquido contendo os antibióticos ampicilina e canamicina nas concentrações de 50  $\mu\text{g}\cdot\text{mL}^{-1}$ . A coexpressão das proteínas recombinantes foi induzida com 0,5  $\text{mmol}\cdot\text{L}^{-1}$  IPTG por 5 horas a 25 °C e

200 rpm. O meio foi centrifugado e as células coletadas. O *pellet* foi resuspendido e lisado em tampão PBS pH 7,4, IGEPAL CA-630 0,5% (v/v), DTT 1 mmol.L<sup>-1</sup> e PMSF 0,1 mmol.L<sup>-1</sup> com lisozima a 0,1 mg.mL<sup>-1</sup> e sonicado. O lisado foi centrifugado a 14.000 rpm, 20 min a 4 °C, e o sobrenadante foi transferido para um tubo de centrifuga de 2 mL onde foi novamente centrifugado. O extrato (1 mL) foi incubado com 30 µL de resina glutationa-Sepharose 4B (GE Healthcare), durante 60 minutos a 4 °C sob agitação constante. Subsequentemente, as amostras foram lavadas três vezes com tampão PBS pH 7,4, IGEPAL CA-630 0,5 % e DTT 1 mmol.L<sup>-1</sup> gelado e as proteínas ligadas à resina foram eluídas com tampão de amostra de SDS/PAGE. As amostras foram fervidas a 95 °C por 5 minutos. As interações foram analisadas em gel de SDS/PAGE.

# **CAPÍTULO 3**

## **Resultados**



### 3. Resultados

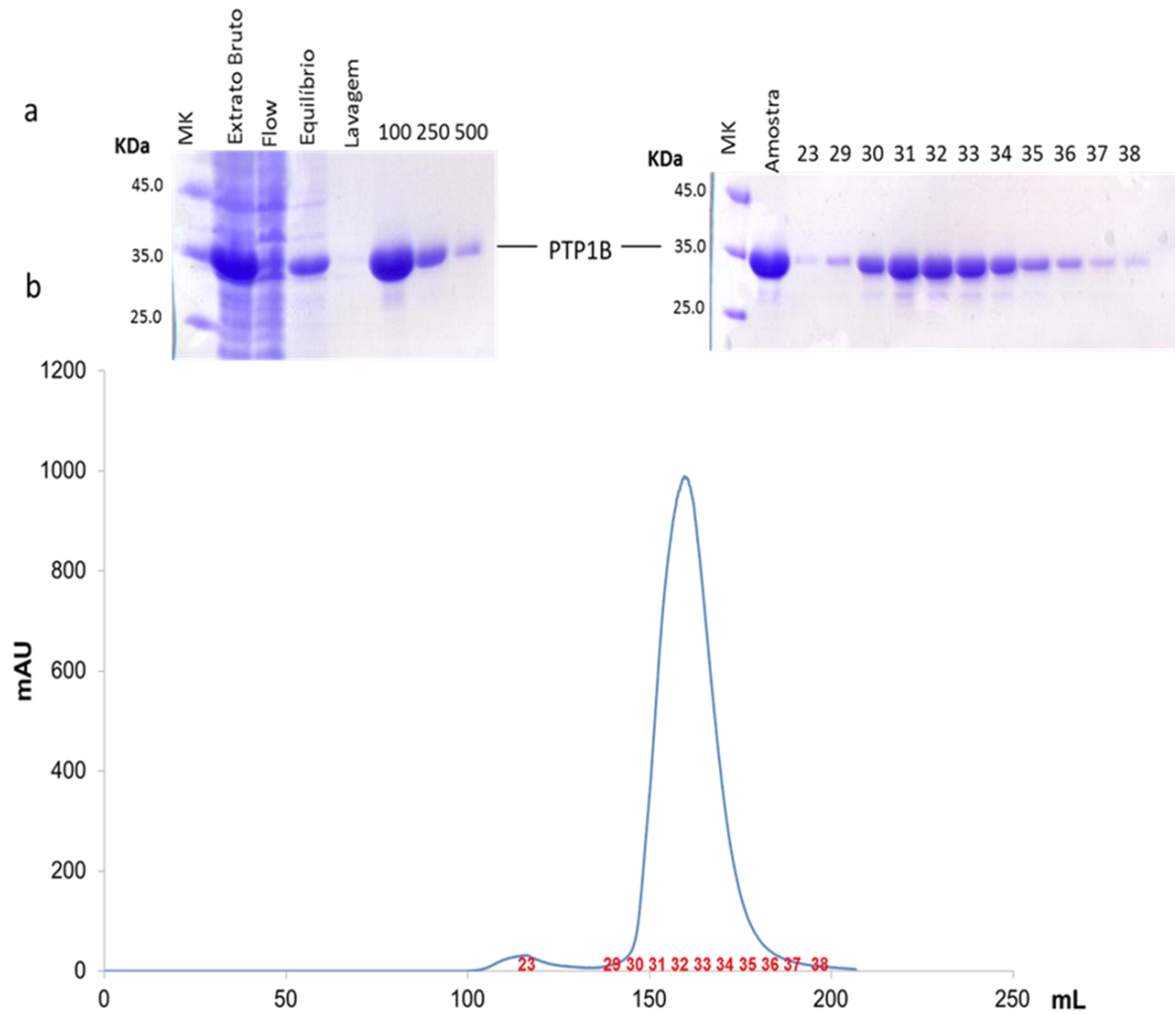
#### 3.1 Estudos com a proteína PTP1B

##### 3.1.1 Ensaios de Inibição Enzimática

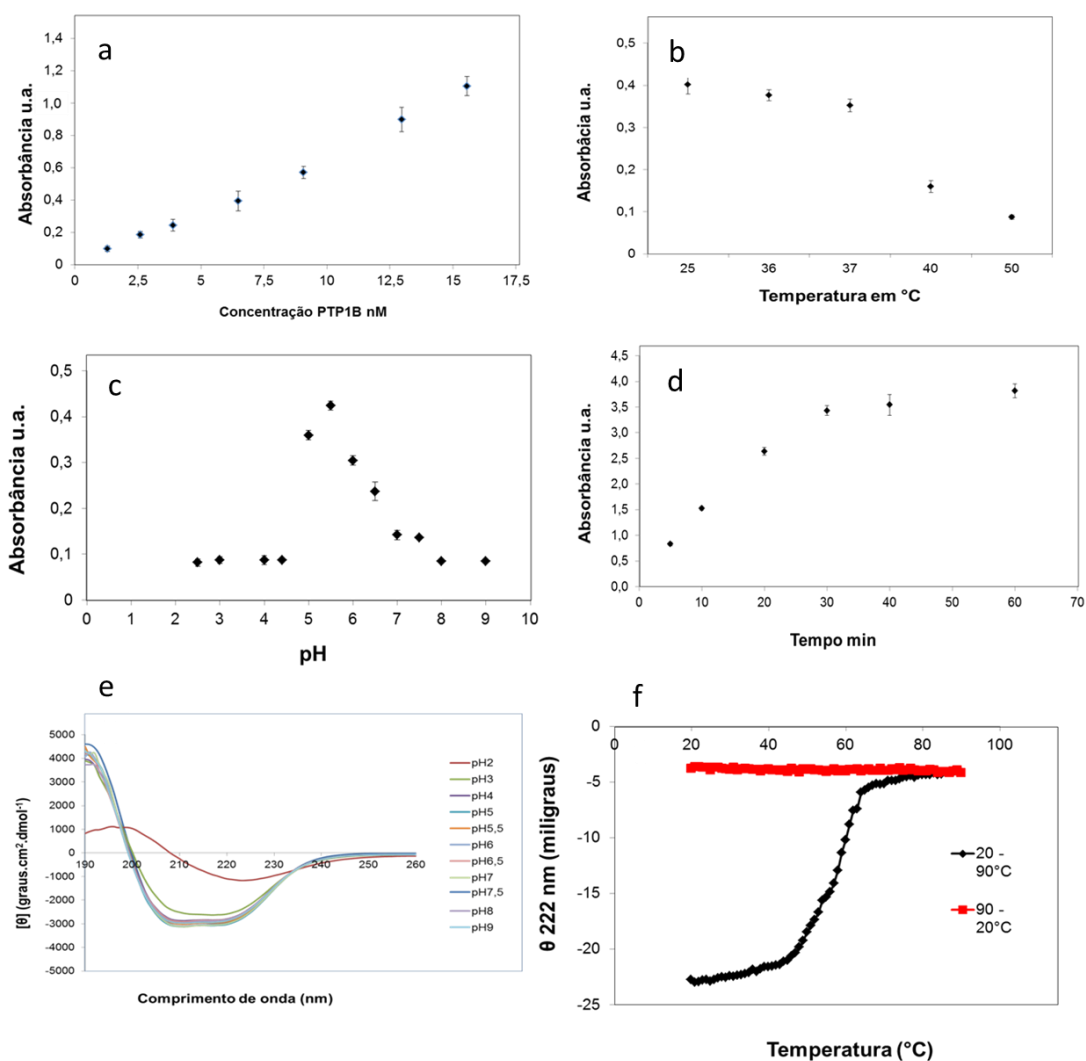
O principal objetivo dos trabalhos com a proteína tirosina fosfatase PTP1B era a obtenção de complexos cristalográficos proteína:ligante, para isso foi necessária a obtenção de proteína pura, corretamente enovelada, ativa e em boa quantidade para os ensaios de triagem de inibidores e de estudos cristalográficos. Aqui serão descritos os principais resultados dos estudos com a proteína PTP1B.

O protocolo de purificação visando a obtenção de proteína para os ensaios foi otimizado para a fosfatase PTP1B baseado no que já existia na literatura (Shi *et. al.*, 2008). A análise do gel (Figura 3.1) revela que a enzima PTP1B está presente em todas as eluições da cromatografia de afinidade por metal imobilizado, principalmente, nas concentrações 100 e 250 mmol.L<sup>-1</sup>. Adicionalmente, também percebemos a presença desta PTP pura em todas as frações da cromatografia de exclusão molecular que seguiram para os estudos posteriores de ensaios enzimáticos para a triagem de inibidores e cristalização de proteínas.

Nos ensaios da PTP1B que serão apresentados a seguir, foi utilizado pNPP (substrato sintético da PTP1B) a 1 mmol.L<sup>-1</sup> (concentração final). Na literatura indica-se que, para atingir o estado estacionário, a concentração do inibidor deve ser de, no mínimo, seis vezes o valor do  $K_M$ , além da relação da concentração enzima:substrato ser de 0,1% (Zabell *et. al.*, 2004). Assim, a concentração de pNPP poderia ser de, no mínimo, 660  $\mu\text{mol.L}^{-1}$ . O  $K_M$  do substrato pNPP é de 110  $\mu\text{mol.L}^{-1}$  para a PTP1B (Hoppe *et. al.*, 1994). Após alguns ensaios preliminares (Figura 3.2), determinou-se que 1 mmol.L<sup>-1</sup> seria uma concentração adequada, já que a absorvância foi satisfatória para a execução dos ensaios. Além disso, verificou-se que o tempo de reação de dez minutos seria suficiente. Também realizaram-se ensaios cinéticos e de dicroísmo circular preliminares para temperatura e pH com o objetivo de otimizar esses parâmetros e assim garantir a qualidade dos ensaios de triagem de inibidores.



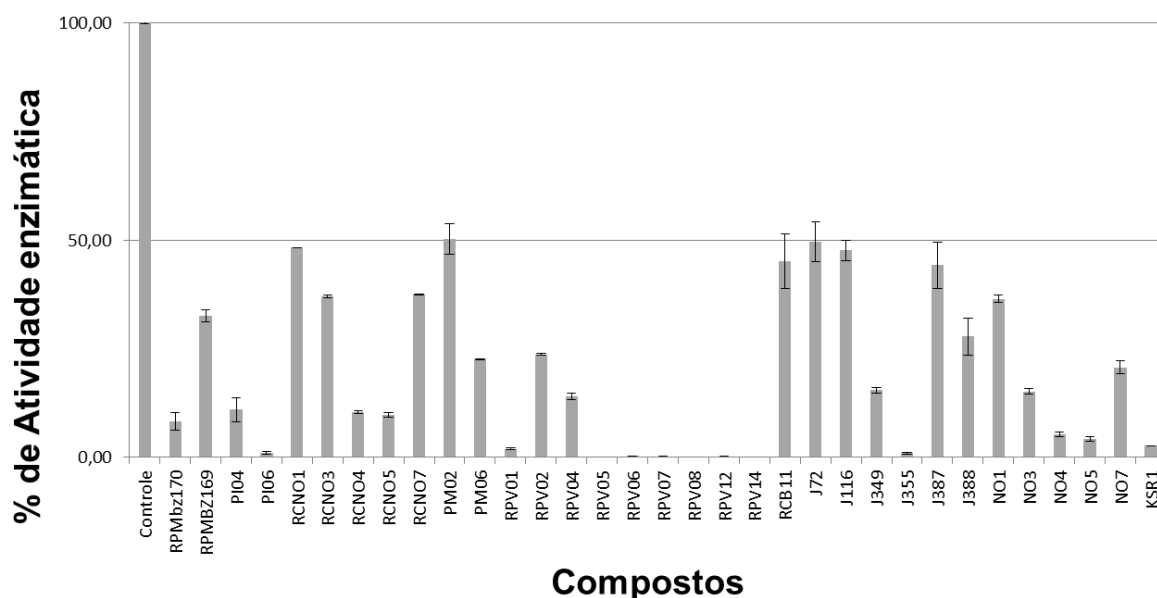
**Figura 3.1:** Purificação da proteína tirosina fosfatase humana PTP1B. **(a)** Os géis de poliacrilamida 12% mostram as eluições da proteína pura, à esquerda IMAC nas concentrações de 100, 250 e 500 mmol.L<sup>-1</sup> de imidazol e à direita número das frações da gel filtração **(b)** O perfil da cromatografia por exclusão de peso molecular em coluna superdex 26/60. A fração 23 corresponde ao *void* e as frações 29 a 38 as eluições da PTP1B pura.



**Figura 3.2:** Ensaios da enzima PTP1B para adequação dos protocolos de inibição com compostos sintetizados pelos grupos de Química Orgânica que participaram do projeto temático. Variação (a) da concentração experimento feito em tampão tris pH 7,5, 100 mmol.L<sup>-1</sup>, pNPP 10 mmol.L<sup>-1</sup>, 10 minutos de reação a 37 °C (b) Temperatura tampão Tris pH 7,5, 100 mmol.L<sup>-1</sup>, pNPP 10 mmol.L<sup>-1</sup>, PTP1B 9,1 nmol.L<sup>-1</sup> tempo de reação 10 minutos (c) pH, reagentes do meio pNPP 10 mmol.L<sup>-1</sup>, PTP1B 9,1 nmol.L<sup>-1</sup> a 37 °C por 10 minutos (d) tempo de reação experimento feito em acetato de sódio pH 5,5, 100 mmol.L<sup>-1</sup>, pNPP 10 mmol.L<sup>-1</sup>, PTP1B 9,1 nmol.L<sup>-1</sup> a 37 °C Espectro de CD para a variação de (e) pH feito a 20 °C, (f) temperatura tampão fosfato de sódio pH 7,5 20 mmol.L<sup>-1</sup>.

Os compostos foram fornecidos para os estudos de triagem e inibição da enzima tirosina fosfatase PTP1B através da colaboração estabelecida com os grupos de Química Orgânica no âmbito do projeto temático: "Biologia Química: novos alvos moleculares naturais e sintéticos contra o câncer. Estudos estruturais, avaliação biológica e modo de ação". Foram cedidos 150 compostos sintéticos e todos foram ensaiados em dose única na concentração de 1 mmol.L<sup>-1</sup> para a proteína PTP1B. Dentre a totalidade de inibidores testados, 22% inibiram a atividade

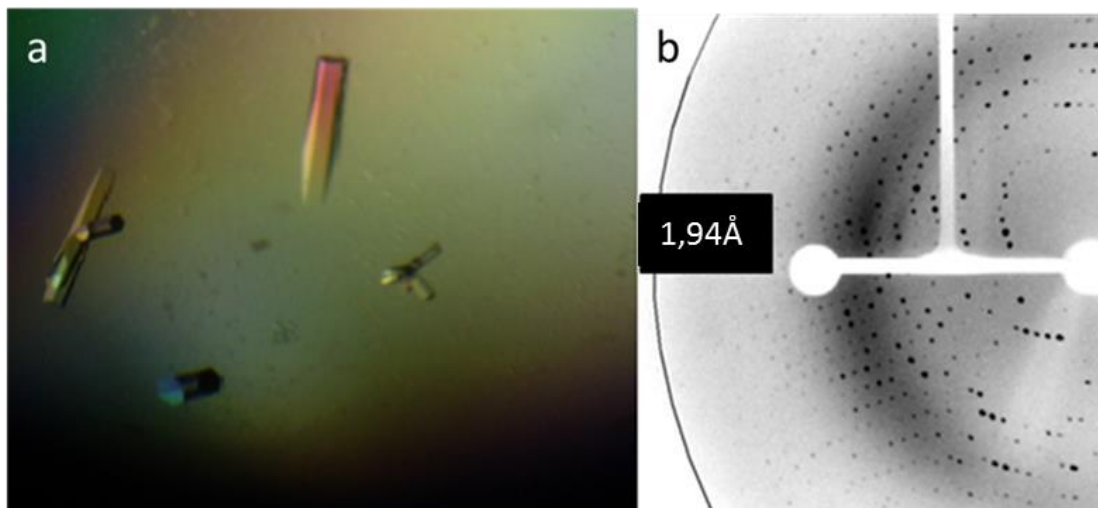
enzimática da PTP1B em 50 % ou mais (Figura 3.3), o que pode ser considerado uma boa taxa de sucesso, uma vez que essas moléculas não foram desenvolvidas a partir de um estudo prévio de desenho racional de inibidores.



**Figura 3.3:** Gráfico dos inibidores da atividade catalítica da PTP1B. Considerou-se como um inibidor a molécula que reduziu a atividade da enzima em 50%.

### 3.1.2 Estudos cristalográficos

A princípio, foram realizados testes de cristalização iniciais como a proteína tirosina fosfatase PTP1B apo para chegar ao protocolo de cristalização ideal, como descrito na seção de Metodologia 2.6.1. A partir da obtenção de bons cristais de proteína (Figura 3.4), várias tentativas de obtenção de complexos proteína:inibidor com os 28 compostos identificados foram feitas, como a cocrystalização e *soaking* a partir dos cristais apo e, também, de cristais providos da cocrystalização.



**Figura 3.4:** Cristais da proteína PTP1B e seu padrão de difração. **(a)** Cristais em formato de bastão sextavado típicos da proteína tirosina fosfatase PTP1B. **(b)** Padrão de difração dos cristais, com círculo desenhado a 1,94 Å.

Os conjuntos de dados foram coletados no laboratório institucional do IQ/Unicamp em difratômetro Bruker Kappa APEX II DUO e na linha de luz MX2 no LNLS. Foram coletados em torno de 60 conjuntos de dados. Porém, os cristais sofrem muito com o processo de *soaking*, o que causou a baixa qualidade dos dados, sendo insuficientes para prosseguir no processamento dos destes. Ao final, 27 estruturas foram obtidas (Tabela 3.1). O limite de resolução dos conjuntos variou de 1.9 a 3.5 Å, sendo que os parâmetros completudeza, redundância e razão sinal/ruído, foram suficientes para realizar a Substituição Molecular e um refinamento inicial, que indicou os resultados expostos na Tabela 3.1. Um conjunto de redução de dados típico de PTP1B foi representado na Tabela 3.2.

Ao longo do tempo, características intrigantes foram observadas próximas ao sítio ativo da PTP1B. Neste local, foi observada, com frequência, densidade eletrônica que aparentava poder ser do ligante. Infelizmente, conforme mostra a Tabela 3.1, constatou-se que havia outras explicações para o fenômeno.

Em oito conjuntos de dados coletados, houve predomínio de estruturas oxidadas na qual há a formação de um intermediário sulfenil-amida. Esse intermediário se dá na presença de espécies reativas de oxigênio e, devido a isso, a proteína se reestrutura. O enxofre da cadeia lateral da cisteína 215 liga-se ao nitrogênio da cadeia principal da serina 216 e o *loop* catalítico é deslocado em direção à C215 (Zhou *et. al.*, 2011; van Montfort *et. al.*, 2003), porém, essa formação

é reversível (Figura 3.5). Quando não há espécies reativas de oxigênio no meio, o *loop* catalítico, além de C215 e S216, permanecem em sua conformação original (van Montfort *et. al.*, 2003). A estrutura 12 (Tabela 3.1), proveniente da tentativa de obtenção do complexo com o ligante PM02, apresentou oxidação de segundo estágio, na qual há dois átomos de oxigênio ligados ao enxofre da cisteína 215. Esse estágio de oxidação da cisteína catalítica é irreversível e indica que a oxidação se deu de modo muito veloz (van Montfort *et. al.*, 2003) (Figura 3.6).

Nos conjuntos 25, 26 e 27, obtidos a partir após o *soaking* com o composto KSR1, há uma densidade eletrônica peculiar próxima aos resíduos catalíticos C215 e R221, provavelmente um indicativo da presença do ligante (Figura 3.7). Todas as tentativas para modelar o ligante foram feitas, porém sua identificação e modelagem não foram possíveis.

Várias estratégias foram adotadas em busca da obtenção de complexos, empregando-se diferentes protocolos de cocristalização e, no caso de *soaking*, variando-se a concentração do inibidor na gota de *soaking* (entre poucos mmol.L<sup>-1</sup> até concentrações supersaturantes) e o tempo de imersão (entre segundos a dias). No entanto, não foi obtido sucesso, provavelmente, devido aos inibidores atuarem de forma promíscua como a série RPV0 que são análogos do produto natural RK-682 (Carneiro *et. al.*, 2015).

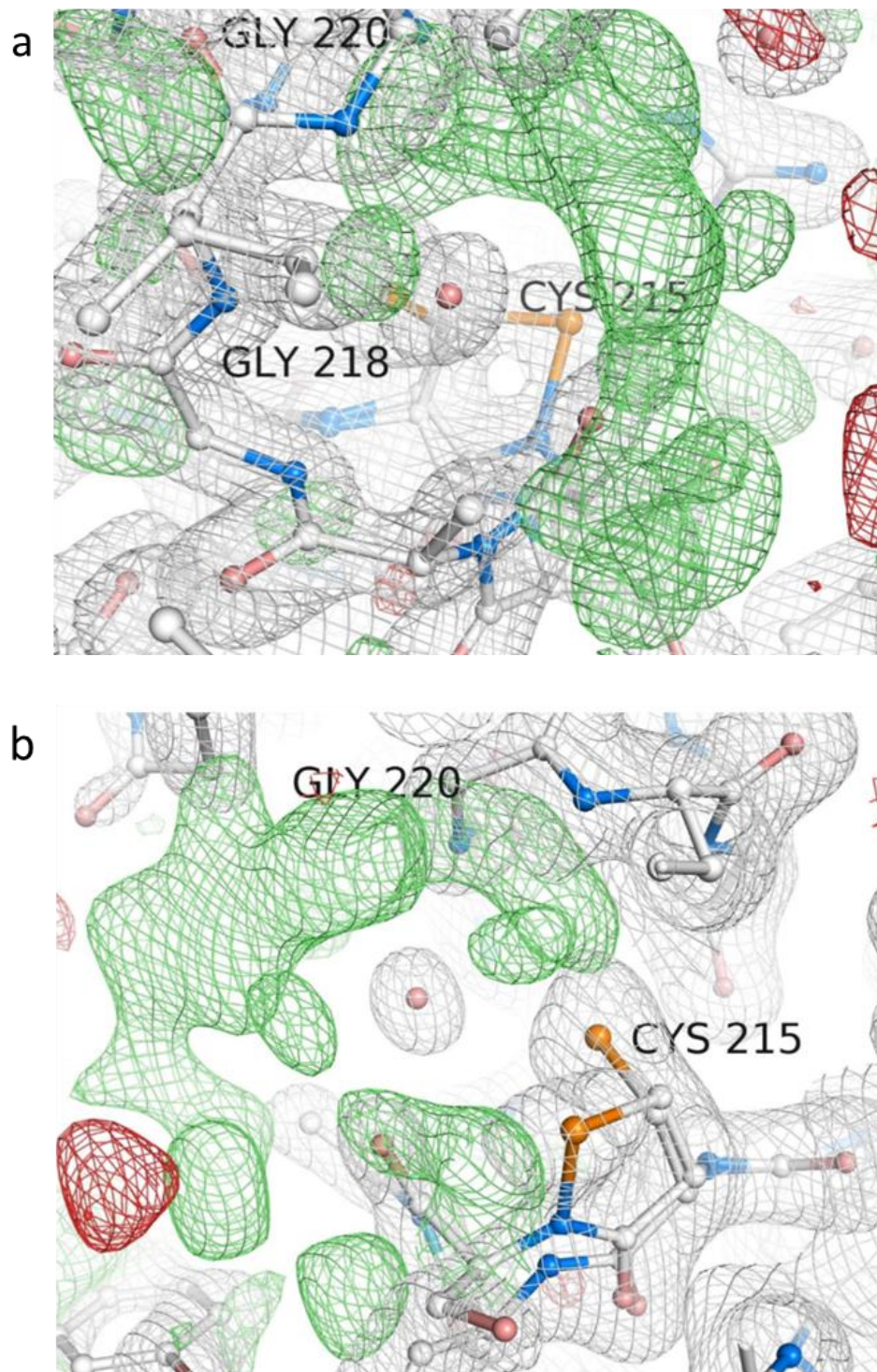
**Tabela 3.1:** Descrição dos conjuntos de cristais coletados para a fosfatase PTP1B. A proteína passou por processos de cocrystalização e os cristais por *soaking* para obter estruturas cristalográficas proteína:inibidor. Esses conjuntos foram coletados em difratômetro APEX II (Bruker) exceto os conjuntos com \* que foram coletados na linha de luz MX2 LNLS. Sem evidências do ligante significa que não foi observada densidade eletrônica correspondente ao ligante. As informações dos compostos 4 a 10 e 15-16 foram previamente obtidas pelo grupo e por isso não constam na Figura 3.1.

#	Inibidor	Método	Resultado	Resolução (Å)
1	APO		Proteína APO	3,5
2	APO		Proteína APO	2,5
3	PI06	Cocrystalização	Sem evidências do ligante	2,8
4	LBCA06	Cocrystalização	Sem evidências do ligante	2,5
5	PIDT	Cocrystalização	Sem evidências do ligante	2,4
6	PIST	Cocrystalização	Sem evidências do ligante	2,8
7	PIST	Cocrystalização	Sem evidências do ligante	2,8
8	LB4A	Cocrystalização	Sem evidências do ligante	3,0
9	LB4A	Cocrystalização	Sem evidências do ligante	3,0
10*	PIST	Cocrystalização e <i>Soaking</i>	Intermediário sulfenil-amida	1,9
11*	PI06	<i>Soaking</i>	Intermediário sulfenil-amida	2,0
12*	PM02	Cocrystalização e <i>Soaking</i>	Segundo estágio de oxidação	2,4
13*	PI06	<i>Soaking</i>	Intermediário sulfenil-amida	1,9
14*	PIDT	<i>Soaking</i>	Intermediário sulfenil-amida	2,4
15*	RPJ285	<i>Soaking</i>	Intermediário sulfenil-amida	2,0
16	FCMSPIR	<i>Soaking</i>	Intermediário sulfenil-amida	2,7
17	PI04	<i>Soaking</i>	Sem evidências do ligante	3,2
18	RCN05	<i>Soaking</i>	Sem evidências do ligante	2,2
19	RCN05	<i>Soaking</i>	Intermediário sulfenil-amida	2,8
20	MBZ169	<i>Soaking</i>	Sem evidências do ligante	3,2
21	MBZ169	<i>Soaking</i>	Sem evidências do ligante	3,2
22	RPV01	Cocrystalização	Sem evidências do ligante	2,0
23	RPV05	Cocrystalização	Sem evidências do ligante	2,4
24	RPJ355	<i>Soaking</i>	Sem evidências do ligante	2,8
25	KSR1	<i>Soaking</i>	Densidade desconhecida próxima a C215	3,0
26	KSR1	<i>Soaking</i>	Densidade desconhecida próxima a C215	2,5
27	KSR1	<i>Soaking</i>	Densidade desconhecida próxima a C215	3,0

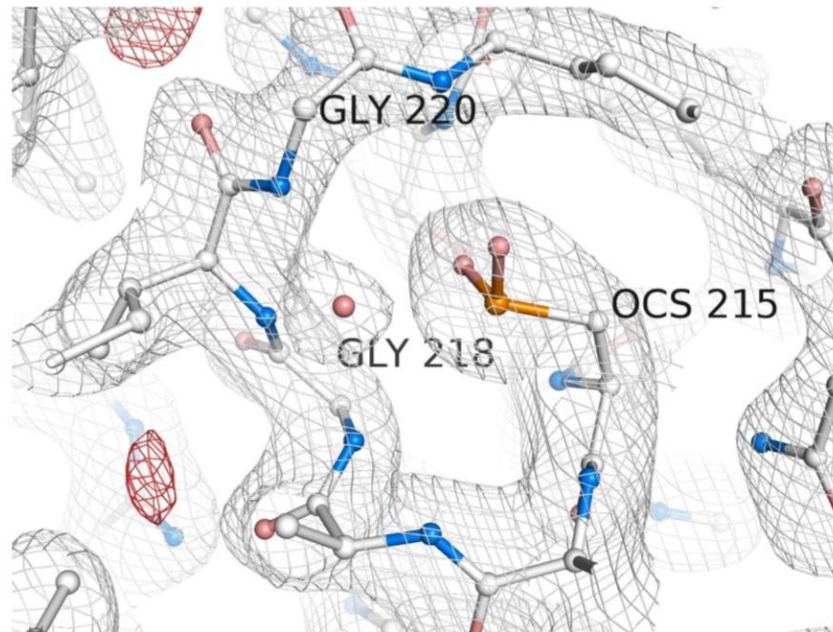
**Tabela 3.2:** Tabela de redução de dados de coleta de um conjunto típico de PTP1B. Dados coletados no difratômetro de Raios X Bruker.

Grupo espacial	P3 <sub>1</sub> 21
<u>Parâmetros de rede</u>	
a (Å)	87,305
b (Å)	87,305
c (Å)	97,540
$\alpha$ (°)	90
$\beta$ (°)	90
$\gamma$ (°)	120
Limite de resolução (Å)	75,61 – 2,60
Última faixa de resolução (Å)	2,69 – 2,60
Mean I	76,9 (15,7)
$R_{int}$	0,4694 (0,6980)
Nº de reflexões totais	13600 (1287)
$\langle I/\sigma(I) \rangle$	10,4 (1,83)
Completeza (%)	99,6 (99,5)
Redundância	9,44 (6,98)

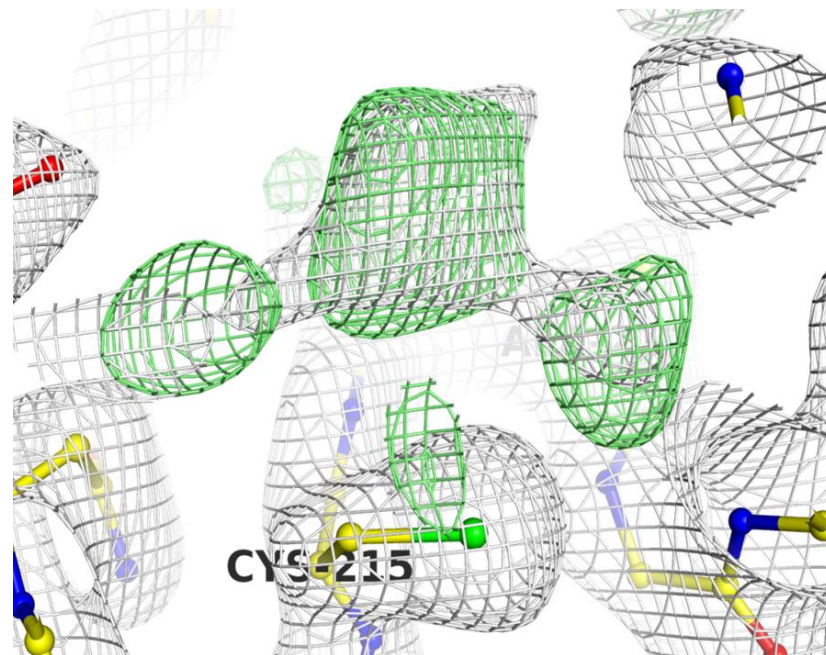




**Figura 3.5:** Mapa de densidade eletrônica da proteína PTP1B na presença do inibidor RPJ285 . As figuras referem-se ao conjunto RPJ285 (no. 15 na Tabela 1:  $R_{free} = 21,4\%$ ,  $R_{factor}=18,2\%$ ) e exemplificam a dificuldade em reconhecer a presença de um possível ligante que, neste caso, está ausente. (a) a figura mostra que o *loop* 216-220 assume dupla conformação na estrutura, conforme ilustra o mapa em verde. (b) nesta estrutura, além de dupla conformação do *loop* em (a), C215 também assume dupla conformação, uma delas correspondendo ao intermediário sulfenil-amida. Mapa de trabalho ( $2F_{obs}-F_{calc}$ ) desenhado a  $1.2 \sigma$  (em cinza) mostrando regiões adequadamente modeladas. Mapa de diferença ( $F_{obs}-F_{calc}$ ) positivo desenhado a  $3.0 \sigma$  (em verde) e  $-3.0 \sigma$  (em vermelho): o primeiro indica locais onde faltam átomos no modelo e o segundo onde há átomos modelados incorretamente. Figura preparada com Pymol (<http://www.pymol.org/>).



**Figura 3.6:** Mapa de densidade eletrônica da proteína PTP1B na presença do inibidor PM02. A figura refere-se ao conjunto PM02 (no. 12 na Tabela 1:  $R_{free} = 26,9\%$ ,  $R_{factor}=20,6\%$ ), em perspectiva similar ao da figura 3.4. Nesta estrutura, observa-se C215 em segundo estágio de oxidação, com o átomo de enxofre ligado a dois átomos de oxigênio. Figura preparada com Pymol.



**Figura 3.7:** Mapa de densidade eletrônica da proteína PTP1B na presença do inibidor KSR1. A figura refere-se ao conjunto KSR1 (no. 6 na Tabela 2:  $R_{free} = 28,9\%$ ,  $R_{factor}=21,4\%$ ) e exemplifica a dificuldade em reconhecer a presença de um possível ligante. Mapa de trabalho ( $2F_{obs}-F_{calc}$ ) desenhado a  $1.2 \sigma$  (em cinza) mostrando regiões adequadamente modeladas. Mapa de diferença ( $F_{obs}-F_{calc}$ ) positivo desenhado a  $2.5 \sigma$  (em verde), que indica locais onde faltam átomos no modelo. Figura preparada com Pymol.

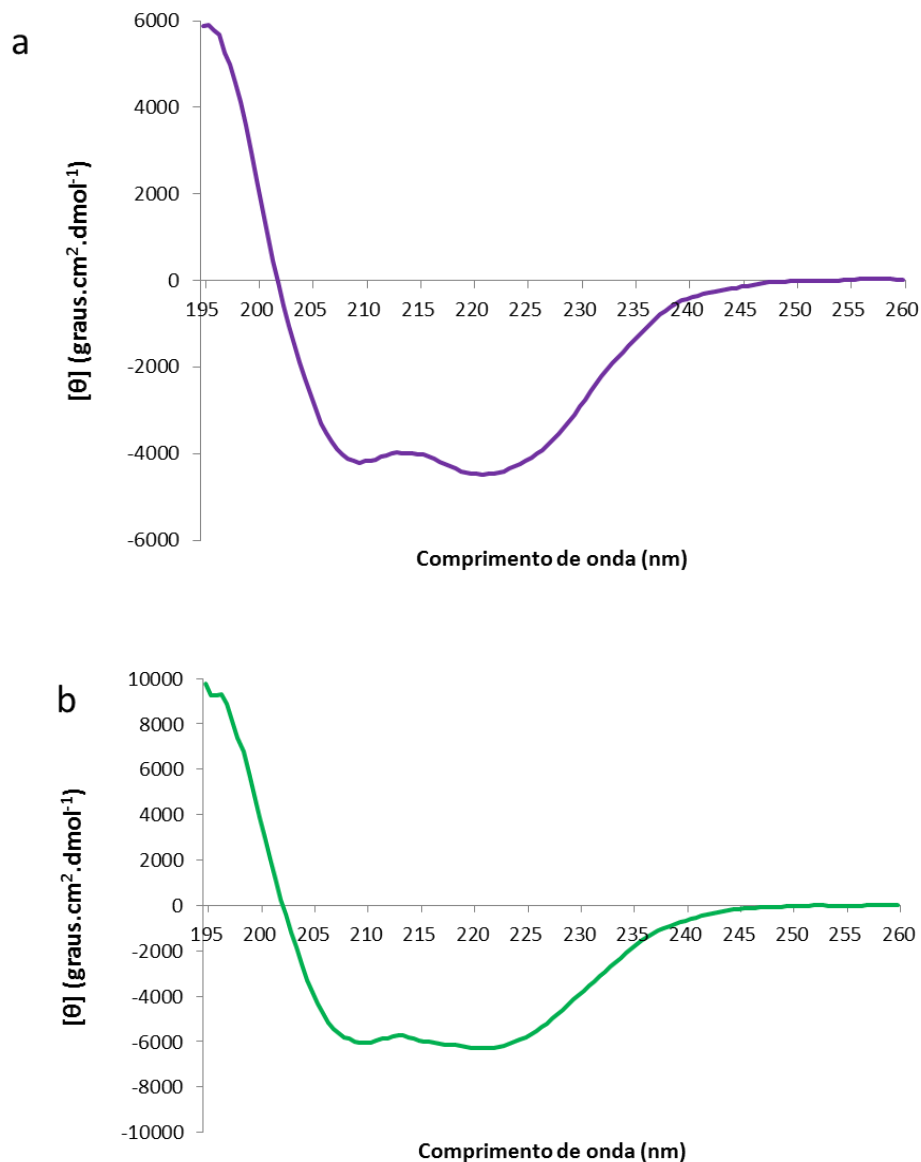
## 3.2 Estudos com os Complexos de PP2Ac

### 3.2.1 Expressão Purificação e Cristalização

O principal objetivo dos trabalhos com os complexos de PP2Ac, heterodímero PP2Ac $\alpha$ 2: $\alpha$ 4 $\Delta$ 222 e heterotrímero PP2Ac $\alpha$ 2: $\alpha$ 4 $\Delta$ 222:TIPRL $\Delta$ N, era a obtenção de amostras puras e bem enoveladas para os estudos cristalográficos. Assim, os complexos foram purificados por IMAC e cromatografia por exclusão molecular, as amostras foram analisadas por DC e DLS e encaminhadas para testes iniciais de cristalização de proteínas.

A migração no gel de poliacrilamida 12% mostra as bandas correspondentes às proteínas que compõem os complexos. Os complexos foram encaminhados para estudos posteriores. A análise de DLS indicou uma polidispersidade de 14% na amostra do heterodímero indicando que esse resultado é aceitável para a cristalização, essa amostra foi encaminhada para o ROBOLAB, junto com o complexo heterotrímero.

As curvas de dicroísmo circular indicaram um predomínio de estruturas em  $\alpha$ -hélices com pico negativos em 222 e 208 nm e pico positivo entre 200 e 190 nm para ambos os complexos (Figura 3.8). Isso já era esperado, uma vez que a estrutura cristalográfica (código PDB 3QC1) demonstrou que a  $\alpha$ 4 é constituída de uma porção N-terminal em TPR (tetratricopeptídeo), composta por sete  $\alpha$ -hélices antiparalelas (Yang *et. al.*, 2007; Jiang *et. al.*, 2013) que interagem com a PP2Ac, e uma outra região intrinsecamente desenovelada no C-terminal (Smetana *et. al.*, 2006). Além disso, a estrutura cristalográfica da PP2Ac $\alpha$  (código PDB 2IAE) (Jiang *et. al.*, 2013; Cho e Xu 2007) mostra que a proteína possui diversos trechos em  $\alpha$ -hélices e algumas folhas- $\beta$ .

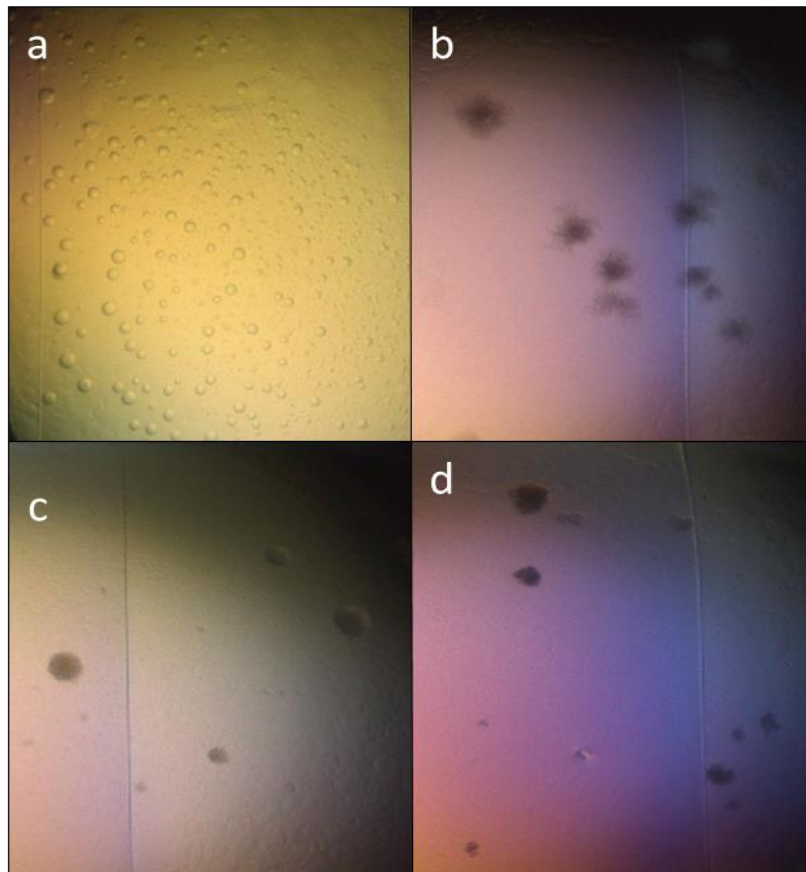


**Figura 3.8:** Curvas de CD para os complexos de PP2A $\alpha$ 2. As curvas apresentam dois picos negativos com mínimos em 222 e 208 nm e um pico positivo na região de 200 a 190 nm. Todos característicos de  $\alpha$ -hélices. **(a)** Complexo heterodímero **(b)** Complexo heterotrímico.

Dando prosseguimento, foram realizados ensaios de cristalização utilizando os kits de cristalização Hampton Crystal Screen I e II, Hampton PEG Rx1; Hampton PEG Ion Screen, com complexos nas concentrações de 5, 10, 13,5 e 15 mg.mL<sup>-1</sup>, representado quase 800 gotas de cristalização no método de gota suspensa a 18 °C, para o complexo heterodimérico. E para as amostras do complexo heterotrímico foram realizados os ensaios de cristalização utilizando os

mesmo kits de cristalização da Hampton nas concentrações de 5 e 13,5 mg.mL<sup>-1</sup>; temperatura de 18 °C.

As amostras enviadas para o ROBOLAB estavam nas concentrações de 2,7 e 10 mg.mL<sup>-1</sup>, para o dímero e 10 e 19,4 mg.mL<sup>-1</sup> para o trímero. Foram testadas 600 condições de cristalização para cada amostra. Enquanto não houve indícios de cristalização para o complexo heterotrímero, para o complexo heterodimérico surgiram alguns pequenos sinais de uma possível nucleação (Figura 3.9), no entanto não cresceram cristais. Experimentos de SAXS indicaram que ambos complexos estavam agregados.



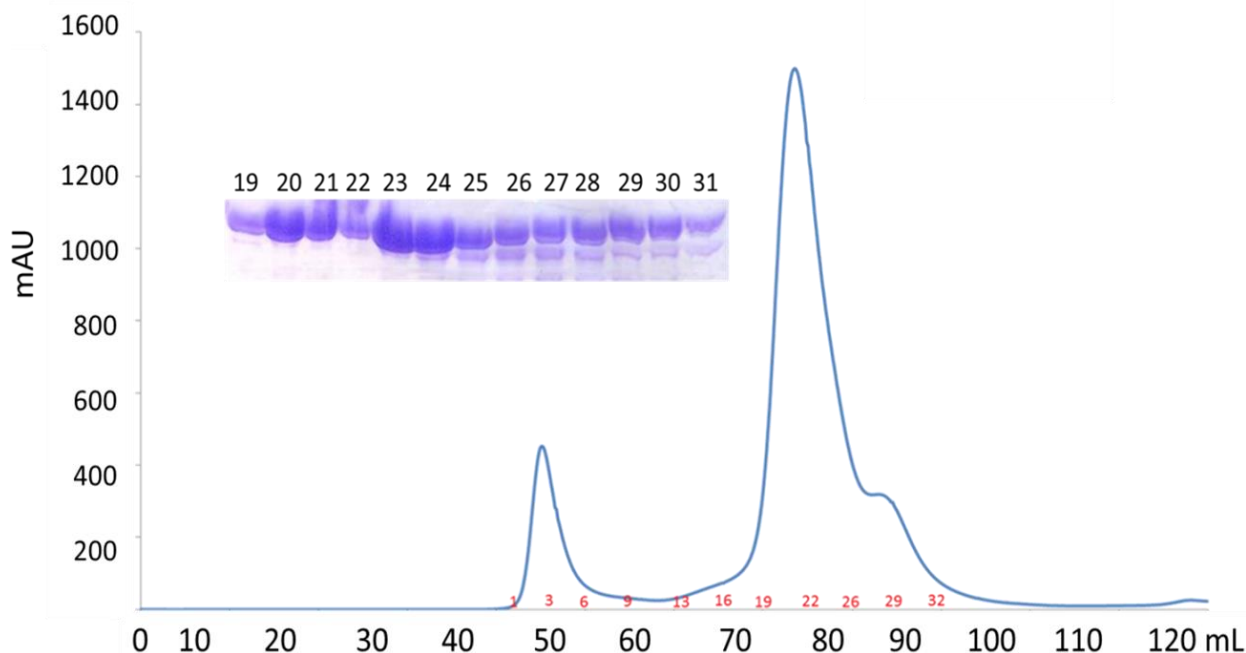
**Figura 3.9:** Indícios de formação de cristais das amostras do complexo heterodímero. (a) Solução PEG 400 50% (v/v); acetato de sódio, pH 4,5, 0,1 mol.L<sup>-1</sup>; sulfato de lítio 0,2 mol.L<sup>-1</sup>. (b) PEG 8000 20% (v/v) CHES pH 9,5, 0,1 mol.L<sup>-1</sup>. (c) Bicine pH 8,5, 0,1 mol.L<sup>-1</sup>; PEG 6000 20% (v/v). (d) PEG 6000 10% (v/v); HEPES pH 6,5, 0,1 mol.L<sup>-1</sup>.

### 3.3 Estudos com a proteína TIPRL

#### 3.3.1 Expressão, purificação e cristalização

O protocolo de expressão foi desenvolvido inicialmente pela coorientadora deste trabalho. O protocolo de purificação de ambas as TIPRL *Full* e  $\Delta N$  foi aperfeiçoado usando a técnica de SAXS, na qual foram realizadas medidas da proteína em diversos tampões, como descrito em Metodologia 2.7.1. As curvas de espalhamento de Raios X a baixo ângulo indicaram que a proteína estava agregada em tampão fosfato, o que também foi observado no dicróismo circular, e melhor estruturada em tampão Tris. Com esse resultado, as concentrações de cloreto de sódio e do agente redutor,  $\beta$ -mercaptoetanol, puderam ser refinadas até a obtenção da solução final de Tris pH 7,5, 20 mmol.L<sup>-1</sup>; NaCl 25 mmol.L<sup>-1</sup>,  $\beta$ -mercaptoetanol 10 mmol.L<sup>-1</sup>, a qual guiou o desenvolvimento dos protocolos finais de purificação e cristalização. A cromatografia de exclusão molecular foi utilizada para separar a proteína TIPRL $\Delta N$  da protease TEV, como pode-se observar na Figura 3.10. Nas últimas frações 25 a 31, há uma mistura entre TIPRL $\Delta N$  e TEV, essas frações foram descartadas para os estudos subsequentes. O rendimento médio foi de 30 mg de proteína pura e estável por litro de cultura e segundo os dados de DLS, com baixa polidispersidade (6,8%). Essa etapa de otimização do protocolo de purificação foi fundamental, pois permitiu a continuidade dos estudos cristalográficos com a TIPRL $\Delta N$ .

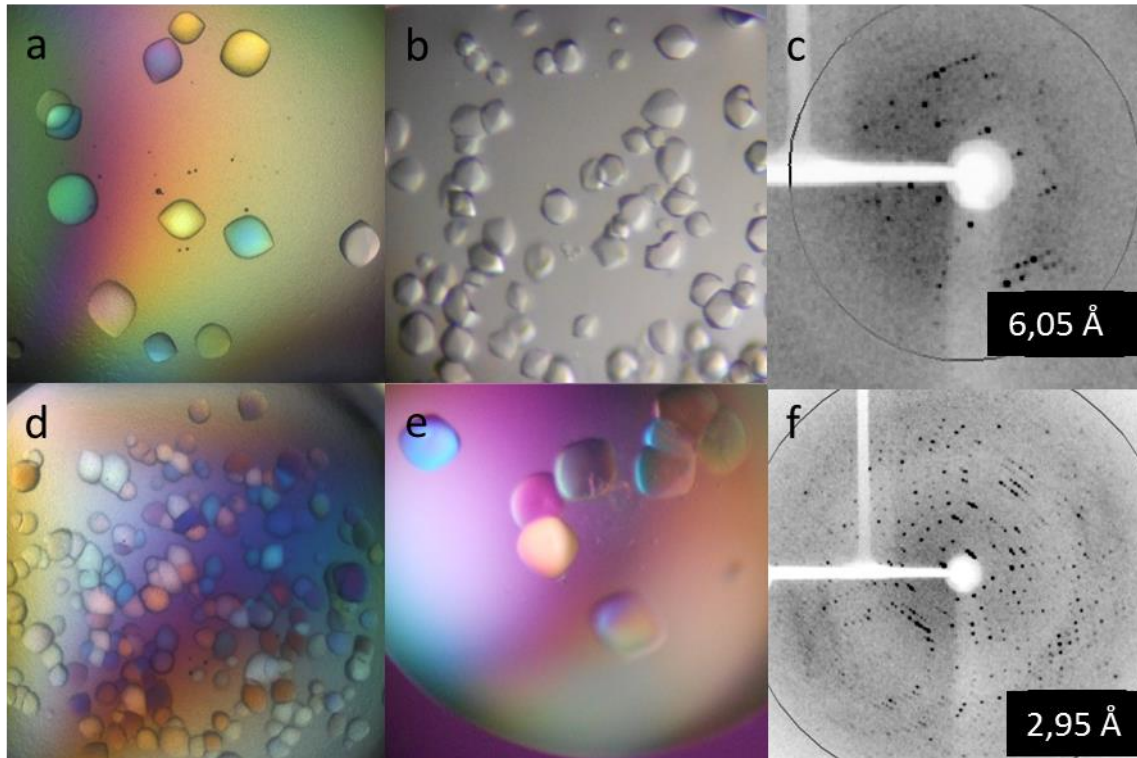
Com a proteína pura e potencialmente monodispersa, foram iniciados os estudos de cristalização. Foram testadas mais de 5.000 condições de cristalização usando kits de cristalização comerciais, variações em tampões, concentração de amostra, tipo de gota, proporções e temperaturas. A alta solubilidade da proteína impediu a sua cristalização no intervalo de concentração habitual de 5-10 mg.mL<sup>-1</sup>, e aumentos graduais na concentração da amostra levaram à identificação da zona de nucleação desta proteína, entre 70-100 mg.mL<sup>-1</sup>. Sinais de cristalização foram observados nas gotas feitas no ROBOLAB (Laboratório Nacional de Biociências) (Tabela 3.2). Essas condições foram reprodutíveis (Figura 3.11).



**Figura 3.10:** Perfil de purificação da TIPRL $\Delta$ N por exclusão de peso molecular em coluna Superdex 200 16/60. Cromatograma da purificação da TIPRL $\Delta$ N. Os números em vermelho indicam as frações da eluição da coluna. Os valores de 1 a 6 representam o *Void*. A proteína pura foi eluída nas as frações 19 até a 32. Em destaque, é apresentado o gel de SDS/PAGE das frações 19 a 31.

**Tabela 3.3:** Condições de cristalização para a proteína TIPRL $\Delta$ N obtidas no ROBOLAB, nas quais houve sinais de cristalização nas concentrações entre 70 a 100 mg.mL<sup>-1</sup> de proteína. A coluna “Tempo” refere-se ao tempo necessário para o crescimento dos cristais nas respectivas condições.

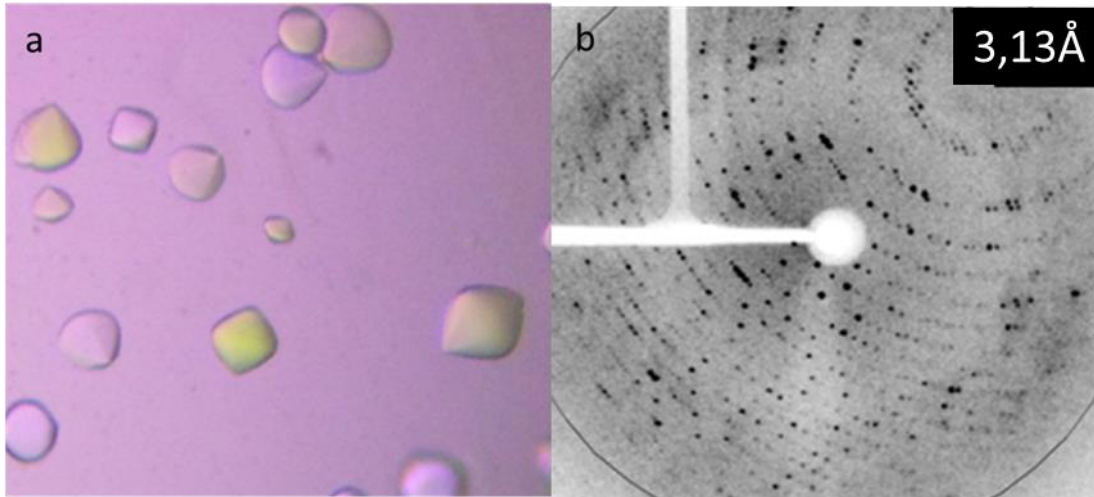
Kit	Condições	Tempo
Wizard B4	Fosfato de sódio di-potássio 0,1 mol.L <sup>-1</sup> pH 6,2 ; Cloreto de sódio 2,5 mol.L <sup>-1</sup>	10 dias
JCSG G12	Bis-Tris 0,1 mol.L <sup>-1</sup> pH 5,5; Cloreto de sódio 3 mol.L <sup>-1</sup>	20 dias
Precipitant F2	Citrato tri-amônio 0,8 mol.L <sup>-1</sup> pH 8,5 ; PEG 8000 2%	2 meses
Precipitant A5	Acetato de sódio 0,1 mol.L <sup>-1</sup> pH 5,5; Cloreto de magnésio 0,1 mol.L <sup>-1</sup> ; PEG 400 2%; Cloreto de sódio 3,9 mol.L <sup>-1</sup>	2 meses
Precipitant E12	Fosfato de sódio di-potássio 1 mol.L <sup>-1</sup> pH 7,5; PEG 4000 0,5%	2 meses
JCSG E8 e Wizard A9	Acetato de sódio 0,1 M pH 4,5 ; Fosfato di-amônio 1 mol.L <sup>-1</sup>	2 meses
Wizard G9	Fosfato de sódio di-amônio 1 mol.L <sup>-1</sup> ; Citrato de sódio 0,1 mol.L <sup>-1</sup> pH 5,5; Cloreto de sódio 0,2 mol.L <sup>-1</sup>	2 meses
Wizard H9	Sulfato de amônio 1,26 mol.L <sup>-1</sup> ; MES 0,1 mol.L <sup>-1</sup> pH 6,0	2 meses
Wizard C10	Fosfato de sódio di-amônio 1 mol.L <sup>-1</sup> ; Imidazol 0,1 mol.L <sup>-1</sup> pH 8,0	2 meses
Wizard E10	Fosfato de sódio di-amônio 1 mol.L <sup>-1</sup> ; Tris 0,1 M pH 8,5	2 meses



**Figura 3.11:** Cristais de TIPRLΔN crescidos no experimento de cristalização do ROBOLAB e reproduzido no laboratório. (a) e (b) cristais nascidos na condição JCSG G12; (d) e (e) cristais na condição Wizard B4; (a) e (d) são imagens das placas de cristalização originais do ROBOLAB; (b) e (e) são os cristais reproduzidos no laboratório; (c) e (f) são padrões de difração dos cristais que comprovam ser de proteína.

Um conjunto de dados de difração de Raios X foi coletado para um cristal de TIPRLΔN crescido em tampão Bis-Tris. Análises com o programa *Cell Now* (Sheldrick 2008) confirmaram que se tratava de um caso de geminação com múltiplos domínios cristalinos. Já os cristais nascidos na condição com o tampão fosfato apresentavam formato esférico e sem faces bem definidas, portanto buscaram-se alternativas para solucionar esses problemas. A primeira foi modificar um pouco a condição de cristalização. Testamos o tampão SPG (*mix* de Fosfato de sódio monobásico; ácido succínico e glicina) em concentrações e pHs distintos. Os cristais crescidos em tampão SPG foram visualmente melhores por apresentarem formas e faces e evidentes semelhantes à de um diamante lapidado (Figura 3.12).

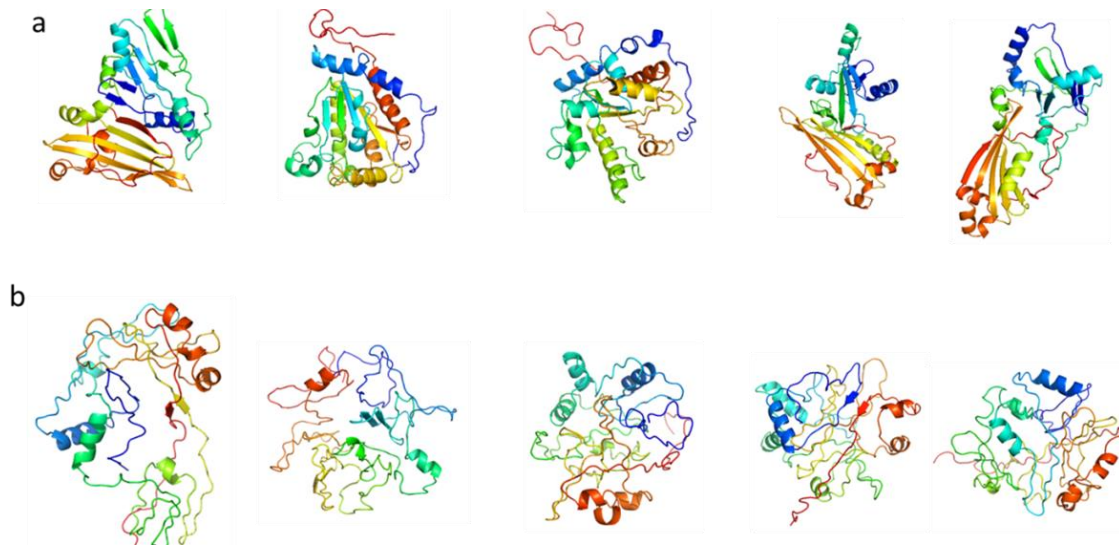




**Figura 3.12:** Cristais de TIPRL $\Delta$ N crescidos em solução de SPG e seu padrão de difração (a) cristais com forma de diamante (b) o padrão de difração dos cristais crescidos em tampão SPG difratando até 3.13 Å, dado coletado no difratômetro Bruker.

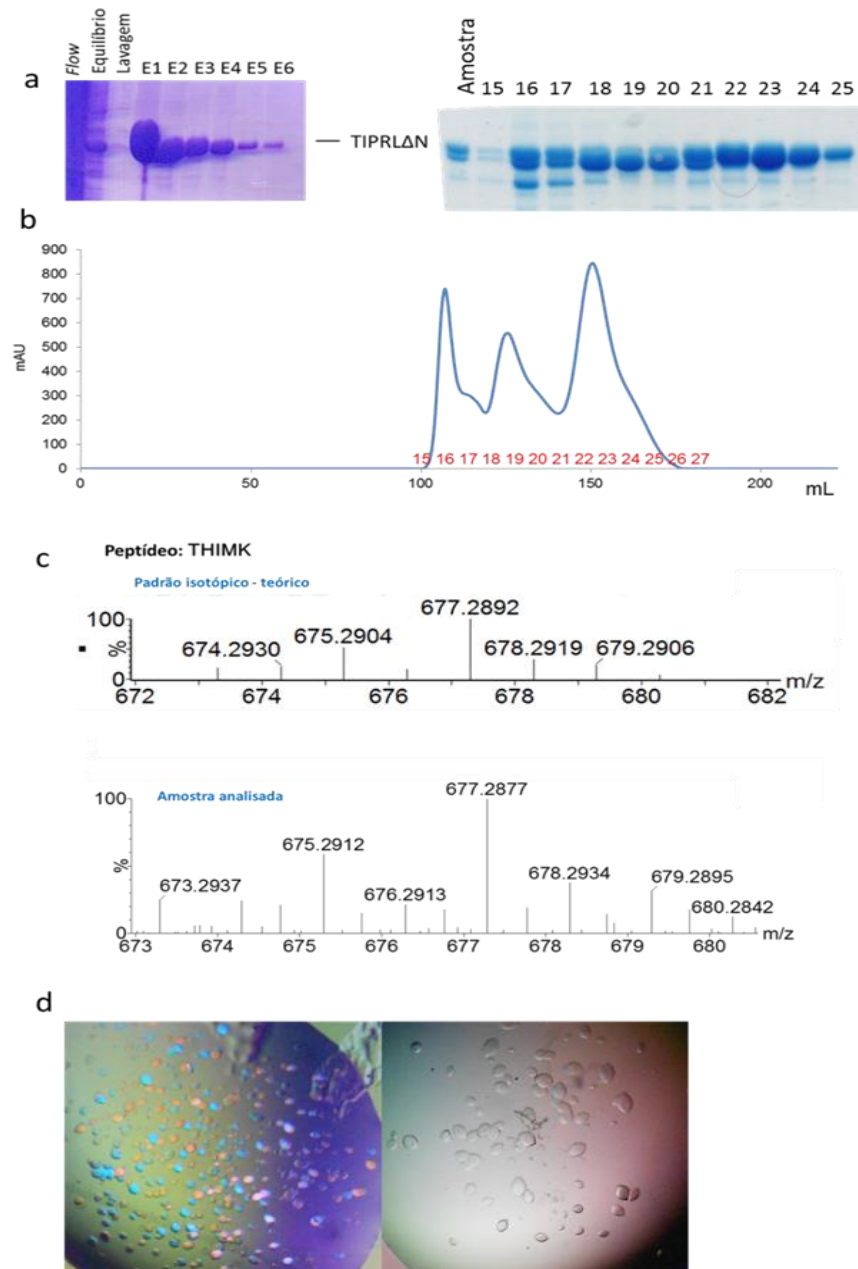
Alguns cristais apo foram coletados com limite de resolução de 2,8 Å e muitas tentativas de Substituição Molecular foram realizadas através de modelos gerados por programas de modelagem molecular como I-TASSER e ROBETTA (Zhang 2008; Roy *et. al.*, 2010; Rohl *et. al.*, 2004). Porém, essa estratégia não funcionou devido à falta de homologia estrutural com outras proteínas depositadas nos bancos de dados. A Figura 3.13 mostra alguns modelos computacionais gerados pelos programas de modelagem. Não há um padrão claro nos modelos, principalmente os modelos gerados pelo servidor I-TASSER que apresentam poucos elementos de estrutura secundária.

Devido à falta de sucesso, partiu-se para a produção da TIPRL $\Delta$ N com seleno-metionina (Se-Met). Essa técnica consiste em crescer as bactérias transformadas em meio mínimo M9 e a metionina fornecida possui um átomo de selênio no lugar do enxofre. Esse átomo pode causar dispersão anômala do cristal (Yang *et. al.*, 1990a; Yang *et. al.*, 1990b; Uchida *et. al.*, 2013). A proteína contendo Se-Met foi cristalizada em condições normais (tampões SPG ou fosfato 0,1 mol.L<sup>-1</sup> pH 5,5 a 6,7 e NaCl 2-3 mol.L<sup>-1</sup>) como descrito em Metodologia seção 2.6.1.



**Figura 3.13:** Modelos gerados computacionalmente para a Substituição Molecular. (a) Modelos gerados no ROBETTA. (b) Modelos gerados no I-TASSER. Figuras em *cartoon* geradas no programa PyMol, em azul a porção N-terminal e em vermelho a porção C-terminal.

O rendimento da proteína purificada se manteve em  $30 \text{ mg.mL}^{-1}$ , o protocolo de purificação se manteve e a solubilidade da amostra não foi alterada, sendo assim pôde-se prosseguir com os estudos. O primeiro experimento com a amostra purificada foi uma análise por espectrometria de massas com o objetivo de garantir que a Se-Met estava incorporada na TIPRL $\Delta$ N, o experimento comprovou que havia átomos de selênio nos peptídeos da proteína (Figura 3.14). Para cristalizar, a proteína foi concentrada isoladamente de cada fração da gel filtração e, também, foram misturadas as frações solúveis 19 a 25 (Figura 3.14). Com essas amostras (isoladas e misturadas), foram realizados os ensaios de cristalização de proteínas, nos quais houve crescimento de cristais nas frações 19, 22, 23 e 25 e na amostra que reunia todas as frações.



**Figura 3.14:** Purificação e cristalização da TIPRLΔN marcada com selenometionina. (a) Gel de SDS/PAGE 12% referente à purificação IMAC as eluições E1-E4 foram feitas com 400 mmol.L<sup>-1</sup> de imidazol e E5-E6, com 500 mmol.L<sup>-1</sup>. Todas as frações apresentam alto grau de pureza. (b) Padrão de cromatografia por exclusão molecular e gel de SDS/PAGE 12%: as frações 15-18 foram eluídas no *void* e foram descartadas por estarem agregadas. As frações puras 19-25 e que foram eluídas em picos diferentes da cromatografia foram encaminhadas para a cristalização de proteínas. (c) Espectro do padrão isotópico do peptídeos THIMK da TIPRLΔN injetados no espectrômetro de massas, comparação do padrão isotópico teórico e experimental do peptídeo modificado com Se-Met indicando que houve a incorporação de Se-Met na proteína. (d) Cristais de TIPRLΔN com Se-Met incorporada.

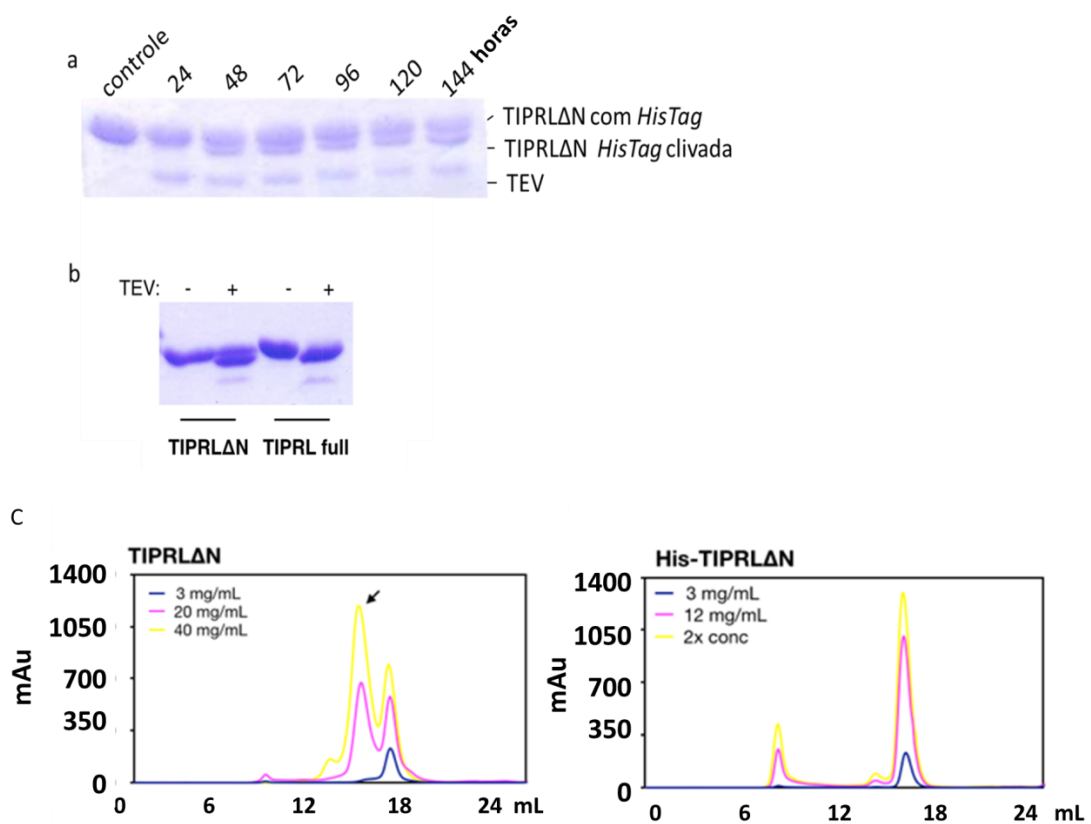
### 3.3.2 Análises do estado oligomérico por SAXS e Gel Filtração

Observou-se um comportamento incomum da TIPRLΔN após a remoção proteolítica da cauda de histidina N-terminal pela enzima TEV, que se correlacionou com a sua capacidade de formar cristais. A amostra submetida a ensaios de cristalização foi, previamente, clivada com a protease TEV, no entanto, a clivagem foi incompleta, resultando geralmente em uma banda dupla na análise de SDS-PAGE (Figura 3.15). Mesmo com tempos prolongados de incubação com TEV, a clivagem da TIPRLΔN permanecia incompleta. Por outro lado, a clivagem da TIPRL *Full* nas mesmas condições era completa, indicando não se tratar de um problema com a enzima proteolítica. Apesar de sua aparente heterogeneidade, a preparação cristalizou, enquanto várias outras preparações que não foram submetidos à clivagem com TEV não formaram cristais. Em contraste, quando TIPRL *Full* foi clivada por TEV sob as mesmas condições, a eficiência da clivagem foi de quase 100%, entretanto esta construção não resultou no crescimento de quaisquer cristais.

Todas as vezes que a amostra de TIPRLΔN clivada com a protease TEV foi concentrada antes de injetar na gel filtração, dois picos eram observados no espectro de UV. Então, as amostras provenientes da gel filtração foram analisadas por SAXS, onde se teve a confirmação de se tratar de duas espécies oligoméricas, dímero e monômero. Além disso, TIPRLΔN, mas não a TIPRL *Full*, foi capaz de dimerizar após remoção da cauda de histidina (Figura 3.15). A análise dos perfis da cromatografia por exclusão de tamanho em diferentes condições permitiu concluir que a dimerização ocorre apenas atendendo a três condições: (i) eliminação dos resíduos N-terminal, (ii) a clivagem da cauda de histidinas por TEV, e (iii) alta concentração proteica, acima de 10 mg.mL<sup>-1</sup>.

Antes de resolver a estrutura cristalina, suspeitou-se que a região N-terminal de TIPRLΔN estava envolvida na interface de dimerização. As análises de *cross-linking*-MS do dímero identificaram uma ligação cruzada de peptídeo intercadeias entre <sup>25</sup>THIMKSADVEK<sup>35</sup>, região N-terminal, que mais tarde foi mapeada na estrutura cristalográfica com uma distância de 19,1 Å entre as K29 das cadeias A e B presentes na unidade assimétrica da cela cristalina.

O Espalhamento de Raios X a Baixos Ângulos (*Small-Angle X-ray Scattering* - SAXS) foi utilizado na análise das construções proteicas em solução. Para cada amostra, foram coletadas séries de imagens de espalhamento com tempos crescentes de exposição, a partir de 3 segundos por imagem. Esta estratégia permitiu verificar que ocorriam alterações estruturais provavelmente relacionadas a danos por radiação. No total, foram coletadas centenas de imagens, que resultaram em centenas de curvas de espalhamento após integração e procedimentos usuais.

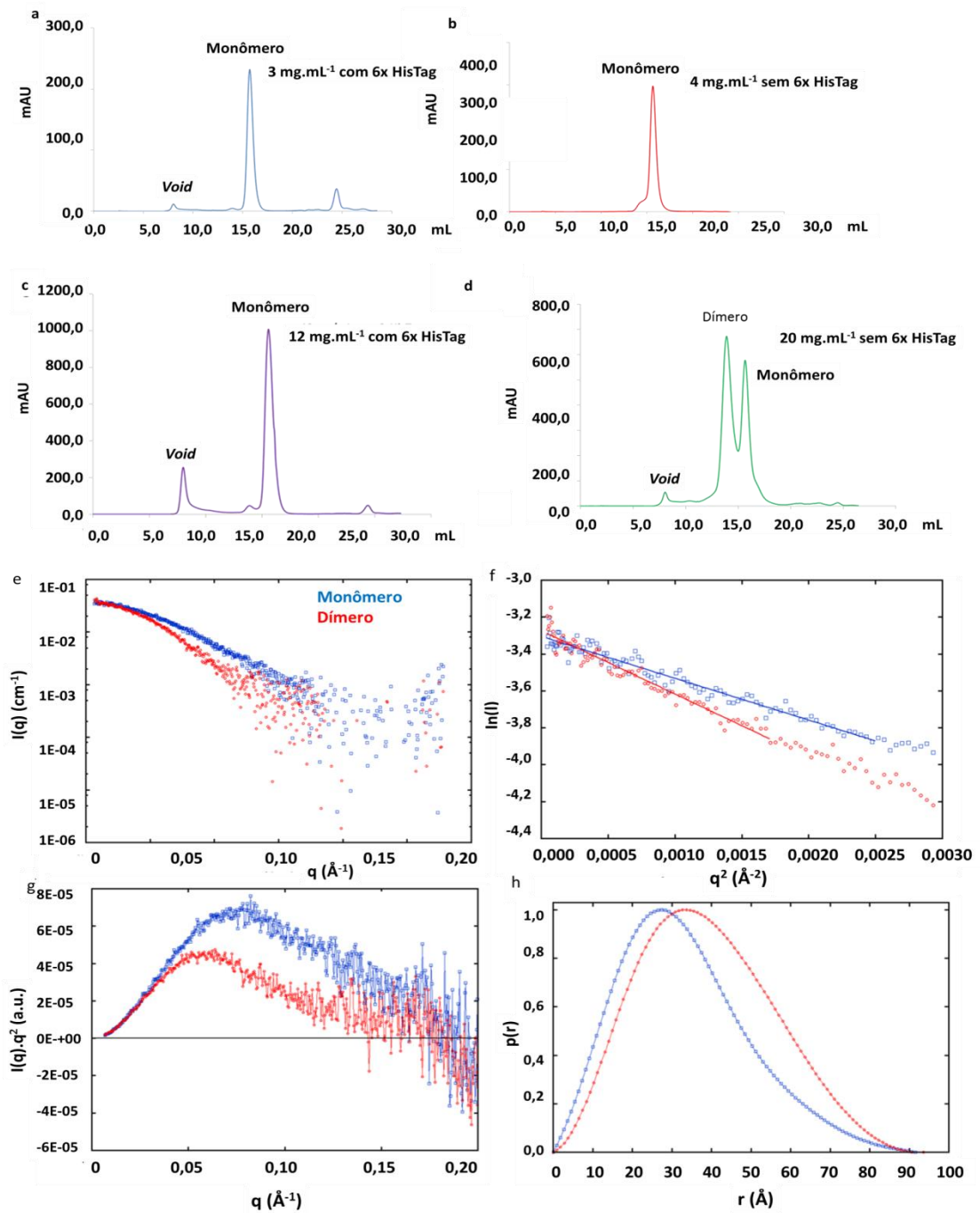


**Figura 3.15:** Dimerização da TIPRLΔN. (a) Monitoramento da clivagem do tag de histidina pela enzima TEV, tempo de 24 a 144 horas. A clivagem é parcial. (b) Gel de SDS/PAGE análise da clivagem da cauda de histidina da TIPRLΔN e TIPRL Full com TEV. Pode-se notar um padrão diferenciado para a construção ΔN. (c) Perfis de gel filtração em coluna Superdex 200 10/30 as amostras em concentrações diferentes e com e sem cauda de histidina. O dímero é formado em altas concentrações e na ausência da cauda de histidina. A seta indica o padrão do dímero na gel filtração.

Além de amostras que porventura estavam agregadas, os resultados indicaram que a TIPRLΔN pode estar presente na forma de monômeros ou dímeros. As duas melhores curvas representativas dessas duas formas são apresentadas na Figura 3.16, e foram obtidas a partir de uma amostra a  $1,4 \text{ mg} \cdot \text{mL}^{-1}$  (monômero) e de

uma amostra a  $0,8 \text{ mg.mL}^{-1}$  (dímero). Uma amostra de  $25 \text{ mg.mL}^{-1}$  foi injetada na gel filtração superdex 10/30 e foi eluído 2 picos, o primeiro representando o dímero e outro representado o monômero. As amostras mais diluídas nesse processo foram analisadas pelo SAXS. Essas amostras vieram na gel filtração, na qual fora Estas são as primeiras curvas de duas séries coletadas e estão isentas de danos por radiação. O baixo tempo de exposição necessário para assegurar esta condição, em conjunto com a baixa concentração das amostras, justifica o ruído relativamente alto observado a maiores ângulos, o que não impediu um processamento estável e confiável dos dados de espalhamento. A massa da partícula espalhadora em solução foi estimada como exposto na metodologia. O método baseado na intensidade normalizada e concentração da amostra levou a estimativas de 37,5 kDa e 65,5 kDa, para monômero e dímero respectivamente. O método independente da concentração confirmou estas estimativas, resultando em 33,3 kDa (monômero) e 67,1 kDa (dímero), próximo ao previsto de 31,3 kDa e 62,6 kDa para o monômero e o dímero respectivamente.

A Figura 3.16 e mostra a regressão linear na região de Guinier, sugerindo soluções monodispersas, enquanto que a Figura 3.16 g aponta para estruturas compactas em solução, típicas de proteínas corretamente enoveladas. A  $p(r)$  mostrada na Figura 3.16 h foi obtida com o programa GNOM, sendo a distância máxima igual a 91,5 Å para o monômero e 93,6 Å para o dímero, com raios de giro respectivamente iguais a 26,1 Å e 29,9 Å. Estas estimativas para o raio de giro concordam com o previsto pela análise de Guinier apresentada na Figura 3.16f (26,1 Å e 32,0 Å para monômero e dímero, respectivamente). Ambos o monômero e o dímero possuem forma globular, levemente alongada, conforme assimetria à esquerda exibida pela  $p(r)$ .



**Figura 3.16:** Cromatografia de exclusão de peso molecular para as amostras de TIPRLΔN com e sem cauda de histidina e curvas de espalhamento de SAXS para a amostra TIPRLΔN sem cauda. Em (a) e (b) são TIPRLΔN em baixa concentração 3 e 4 mg.mL<sup>-1</sup> respectivamente: observa-se um pico único (a) possui cauda de histidina e (b) não possui. Já em (c) e (d) as concentrações são elevadas, 12 e 20 mg.mL<sup>-1</sup>. Em (c) com cauda de histidina há um pico único e um pico protuberante no *void*; já em (d) que possui o *Tag* pode-se observar dois picos indicando a dimerização da proteína em alta concentração sem o *Tag* e um pico mínimo na região do *void*. (e) curva de espalhamento experimental para o monômero e dímero da TIPRLΔN. (f) gráfico de Guinier. (g) gráfico de Kratky. (h) função distribuição de distâncias  $p(r)$ .

### 3.3.3 Coleta de dados

Os cristais crescidos com Se-Met foram enviados para o síncrotron Diamond Light Source, UK, no total de 64 exemplares apo e 16 banhados com cloreto de cério, onde foram coletados de modo remoto em colaboração com o Dr. Humberto Pereira do Instituto de Física/IFSC-USP. Coletaram-se 26 conjuntos de dados apo, todos com boa qualidade.

O Problema das Fases foi resolvido a partir de um conjunto de menor resolução, porém, com excelente sinal anômalo (Tabela 3.4). Foi construído automaticamente um modelo inicial com resolução a 2,66 Å pelo *software* ARP/wARP. O programa construiu e docou corretamente quase a totalidade dos resíduos de aminoácidos que formam a proteína.

O conjunto de dados usados para construir o modelo foi analisado pelo pacote de programas SHELX C/D/E (os sumários dos programas estão logo a abaixo). Este pacote identifica e localiza os átomos que geram o espalhamento anômalo. Diferentes *jobs* de SHELXD foram feitos para cada grupo de espaço: P6, P61, P62 e P63. Para cada grupo, foram buscados 5, 10, 20, 39 e 78 sítios. Com isso, a partir de excelente sinal anômalo conseguiu-se determinar P61 como melhor grupo (dentre os analisados até este ponto), junto com 19 sítios de espalhadores anômalos (de onde foi possível antecipar a presença de 2 moléculas de TIPRLAN na unidade assimétrica).

Na sequência, SHELXE (que calculou as fases a partir dos sítios encontrados pelo SHELXD) foi utilizado com conteúdo de solvente entre 0,25 e 0,75, em intervalos de 0,05. O melhor resultado foi obtido com grupo P65 (enantiomorfo do P61 de entrada), para um conteúdo de solvente de 0,45. Embora fosse sabido que o conteúdo de solvente era maior, 76,2%, assumindo 2 moléculas de TIPRL na unidade assimétrica, programa MATTHEWS\_COEF (Matthews, 1968; Kantardjieff e Rupp, 2003), nesta etapa o algoritmo do SHELXE funcionou melhor com 45% de solvente.



**Tabela 3.4:** Estatísticas de coleta de dados para o conjunto a 2,66 Å. Os valores entre parênteses se referem à última faixa de resolução.

Grupo espacial	P6
<u>Parâmetros de rede</u>	
a (Å)	145,83
b (Å)	145,83
c (Å)	95,93
$\alpha$ (°)	90
$\beta$ (°)	90
$\gamma$ (°)	120
Limite de resolução (Å)	29,29 – 2,69
Última faixa de resolução (Å)	2,72 – 2,66
$R_{merge}$	0,113 (0,786)
$R_{meas}$	0,123 (0,866)
$R_{pim}$	0,047 (0,355)
Nº de reflexões totais	200998 (12787)
Nº de reflexões únicas	33078 (2388)
$\langle I/\sigma(I) \rangle$	12,5 (1,9)
CC(1/2)	0,996 (0,661)
Completeza (%)	99,0 (97,4)
Redundância	6,1 (5,4)
Completeza anômala (%)	90,4 (85,1)
Redundância anômala	2,9 (2,8)
<i>DelAnom correlation between half-sets</i>	0,584 (0,076)
<i>Mid-Slope of Anom Normal Probability</i>	1,314 (-)

Sumário do SHELXC:

<i>Unit cell</i>	a=b145.83 Å c= 95.93 Å $\alpha=\beta=90^\circ$ $\gamma=120^\circ$		
<i>N try</i>	200		
<i>Pointgroup</i>	P6		
<i>Resolution (Å)</i>	29.29 – 2.66		
<i>Nrefl</i>	33077 / 29144		
<i>DF/F</i>	0.102		
<i>dI/sig(dI)</i>	1.276		
<i>Spacegroups</i>	P6 P61 P62 P63		
<i>Dmin</i>	$\langle I/\sigma \rangle$	<i>%comp</i>	$\langle d''/\sigma \rangle$

5.92	27.60	93.30	2.82
4.70	23.40	97.80	1.93
4.10	23.10	98.60	1.54
3.73	18.40	98.90	1.32
3,46	14.90	99.10	1.12
3,26	11.60	99.30	1.01
3,09	7.90	99.20	0.95
2,96	5,50	99.60	0.82
2,84	3.80	99.80	0.83
2,75	2.70	99.70	0.77
2.66	2.00	98.30	0.78

Sumário do SHELXD:

*Running 20 x shelxd\_mp jobs*

*Best spacegroup: P61*

*Best nsites: 19*

*Best CC / weak: 45.47 / 26.55*

Sumário do SHELXE:

*Running 22 x shelxe jobs*

*Best solvent: 0.45*

*Best hand: inverted*

*Best spacegroup: P65*

Os 26 conjuntos de dados coletados foram analisados e reprocessados, onde se encontrou um conjunto a 2,39 Å. Porém, esse conjunto não tinha sinal anômalo apreciável, pois houve mudanças na estratégia de coleta. Então, os dados dos conjuntos foram novamente reprocessados usando o pipeline DIALS (Gildea *et al.*, 2014), que levou a 2,15 Å de resolução. Para manter válido o teste de Rfree, foi copiada a coluna do arquivo .mtz anterior neste novo arquivo. Foram feitas a

**Tabela 3.5:** Estatísticas de coleta e redução de dados para o conjunto a 2,15 Å. Os valores entre parênteses se referem à última faixa de resolução.

<u>Estatísticas de coleta de dados</u>	
Conteúdo de solvente (%)	76,20
Grupo espacial	P65
Parâmetros de rede	
a = b (Å)	145,31
c (Å)	95,95
$\alpha = \beta$ (°)	90
$\gamma$ (°)	120
Resolução (Å)	95,95 – 2,15
Última faixa de resolução (Å)	2,25 - 2,15
$R_{merge}$ (%)	0,119 (1,574)
$R_{meas}$ (%)	0,123 (1,628)
$R_{pim}$ (%)	0,031 (0,413)
Nº Total de reflexões	952365 (125665)
Nº de reflexões únicas	62695 (8301)
$\langle I/\sigma(I) \rangle$	12,1 (2,31)
CC(1/2)	0,995 (0,760)
Completeza (%)	100 (100)
Redundância	15,2 (15,1)
Resolução (Å)	95,95 – 2,15
Completeza Sinal anômalo	100 (100)
Redundância sinal anômalo	7,5 (7,4)
<i>DelAnom correlation between half-sets</i>	-0,079 (0,010)
<i>Mid-Slope of Anom Normal Probability</i>	0,927 (-)
<u>Estatísticas de Refinamento</u>	
Nº de reflexões usadas no $R_{free}$ (4%)	2525
$R_{factor}$ (%)	0,19
$R_{free}$ (%)	0,20
Átomos de proteína	4360
Moléculas de água	233
Átomos inorgânicos	2 Cl <sup>-</sup> , 1 Ni <sup>2+</sup> , 2 PO <sub>4</sub> <sup>2-</sup>
<i>B-factor</i> (Å <sup>2</sup> )	53,338
Desvio da idealidade (RMS):	
Comprimento da ligação (Å)	0,0159
Ângulo da ligação (°)	1,9412
Análise de <i>Ramachandran</i>	
Regiões favoráveis (%)	98
Regiões permitidas (%)	2
Regiões proibidas (%)	0

Substituição Molecular com o modelo a 2,66 Å gerado pelo ARP/wARP e o refinamento neste conjunto com 2,15 Å de resolução. As informações de redução de dados e refinamento podem ser observadas na Tabela 3.5.

### 3.3.4 Estrutura

O cristal pertence ao grupo de espaço hexagonal P65, o que não é muito comum, apenas 1,09% de todas as estruturas cristalográficas depositadas no PDB pertencem a esse grupo, com 2 moléculas na unidade assimétrica e 76,2% de solvente. As moléculas de proteína que formam o cristal estão ligadas de modo covalente por uma ponte dissulfeto entre as cisteínas C87 da cadeia A e da cadeia B. Esta ponte está localizada na porção C-terminal, região oposta à interface de dimerização e foi essencial para o processo de cristalização, pois acomodam de modo ordenado as moléculas dentro do cristal. Foi criado um microambiente oxidativo durante o processo de cristalização, o qual foi viável para a formação dessa ponte dissulfeto mesmo na presença de 10 mmol.L<sup>-1</sup> de β-mercaptoetanol. Já na face oposta da estrutura, o N-terminal forma um dímero entre duas moléculas da proteína. Portanto, o cristal é formado por uma interação covalente no C-terminal e uma interface de dimerização no N-terminal (Figura 3.17).

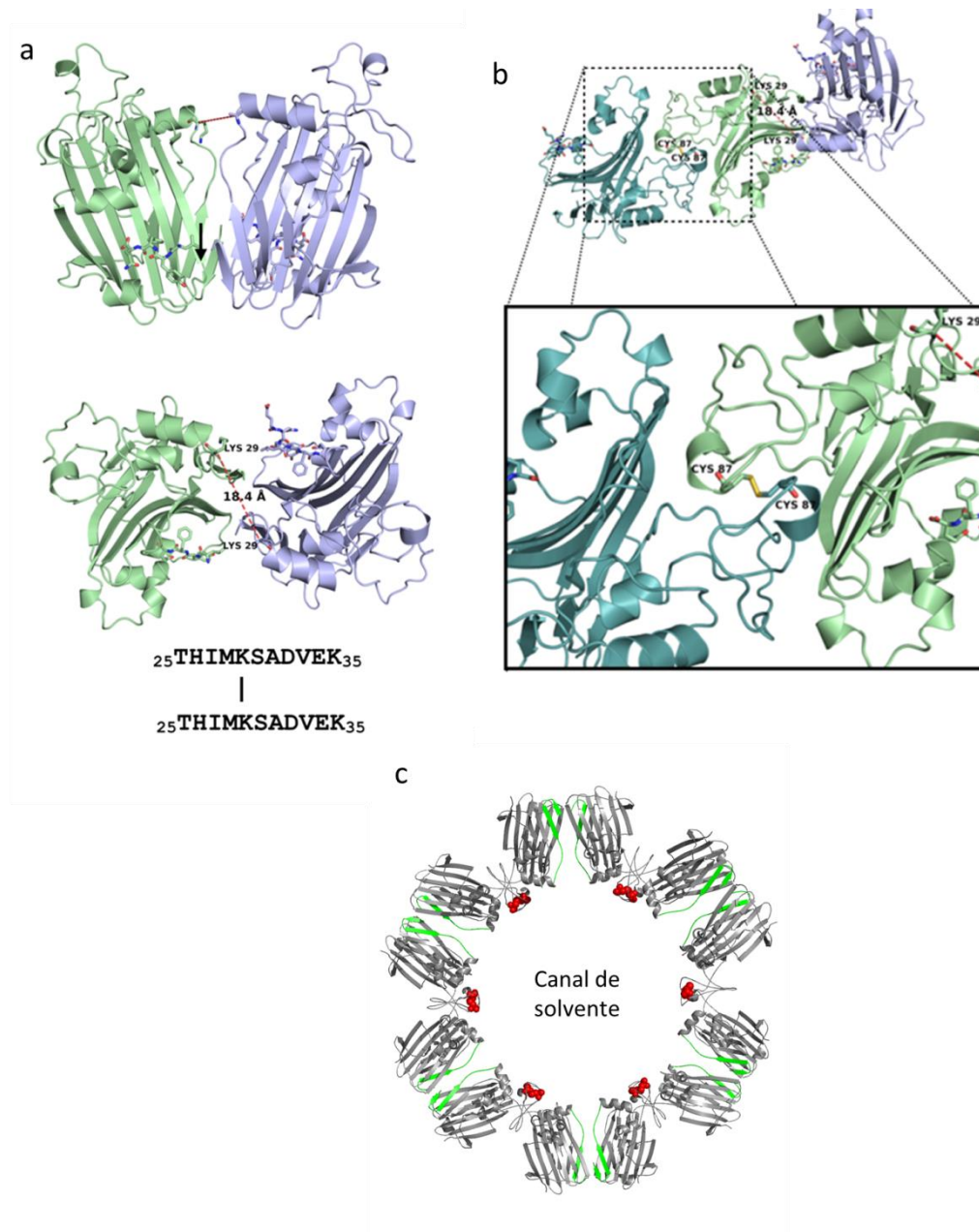
A estrutura da TIPRLΔN é composta por um núcleo de folha-β antiparalela com seis longas fitas de até 12 resíduos, que é coberto por um domínio N, semelhante a uma *lid*, onde uma folha-β menor com cinco fitas está presente, juntamente com duas α-hélices e uma quantidade considerável de resíduos que não adotam nenhum elemento de estrutura secundária conhecido, mas ainda assim são ordenados suficientemente para aparecer no mapa de densidade eletrônica. A *lid* corresponde ao domínio N-terminal, enquanto que o núcleo de folha-β corresponde ao C-terminal que consiste, principalmente, de resíduos conservados. Ambos os domínios estão ligados através de uma longa alça (resíduos 121-148) que contém um segmento curto em conformação de α-hélice, a qual abrange o núcleo de folhas-β e cobre parte do seu lado inferior, formando uma fenda juntamente com um pequeno subdomínio alfa-helicoidal localizado oposto a ele. O domínio C-terminal folha-β apresenta a maior conservação evolutiva, realçando a sua importância

estrutural (Figura 3.18). A estrutura possui três curtas hélices do tipo  $3_{10}$  (hélices em  $\eta_1$ ,  $\eta_2$  e  $\eta_3$  na Figura 3.18), em que uma volta completa envolve 3 resíduos de aminoácidos rotacionados a  $120^\circ$ , relativamente.

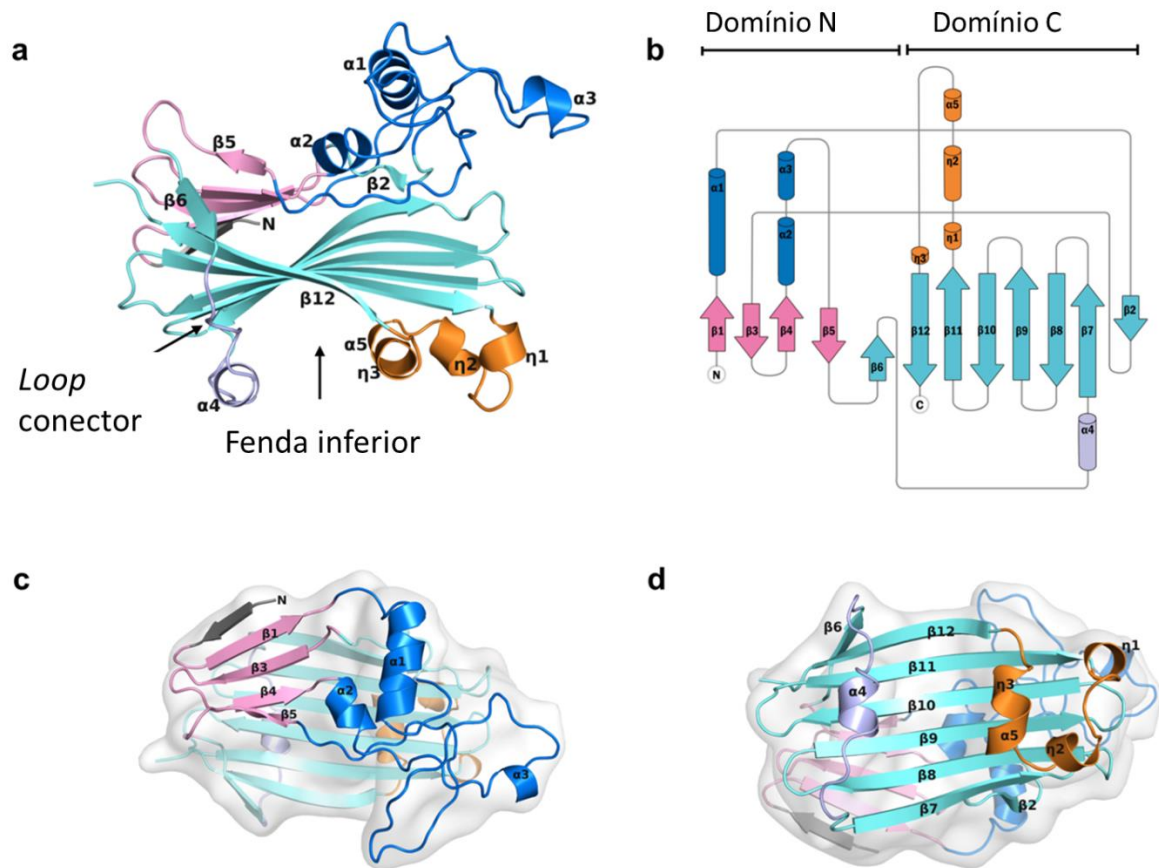
A estrutura foi submetida aos bancos de dados de semelhança estrutural *DALI*, *TOP Search* e *CATH* a fim de encontrar semelhanças estruturais com outras proteínas, as estruturas que mais se assemelharam à TIPRLΔN foram estruturas do tipo barril- $\beta$ , porém o enovelamento da TIPRLΔN não é desse tipo. Portanto, a TIPRLΔN exibe um novo *fold* que ainda não havia sido descrito, o que não ocorria desde 2008, segundo estatísticas do PBD.

A densidade eletrônica dos últimos 22 resíduos C-terminal da TIPRLΔN não foi observada no mapa de densidade eletrônica. Esta região composta pelos últimos resíduos não é conservada, mas abriga um possível local de fosforilação validado experimentalmente, uma vez que a serina S239 foi indicada como possível local de fosforilação (Matsuoka *et. al.*, 2007). A S239 seria um substrato da enzima quinase ATM/ATR, a qual adicionaria um fosfato com a provável finalidade de regulador do sítio de ligação com a PP2Ac, tendo em vista que o resíduo S239 está localizado próximo à cavidade de interação com o C-terminal da PP2Ac.

O mapa de diferença  $|F_{\text{obs}} - F_{\text{calc}}|$  apresentou uma região não explicada que precisava de átomos nas proximidades do resíduo de aminoácido I136. A princípio, foi modelado uma dupla conformação do *loop* ou a cauda de histidina que não foi clivada, no entanto sem sucesso. Tentou-se, então, modelar os quatro últimos resíduos de aminoácidos do sítio de clivagem da proteína TEV, LYFQ (Figura 3.19), o qual se encaixou perfeitamente na estrutura e satisfaz a diferença no mapa de densidade eletrônica. Após a clivagem com a TEV, o peptídeo (MGHHHHHENLYFQ) gerado estava presente na solução e mesmo depois da etapa de gel filtração, permaneceu ligado à TIPRLΔN, indicando uma forte afinidade pela proteína. Porém, há uma diferença nas densidades eletrônicas dos peptídeos provindos da cauda de histidina. Na molécula A, a densidade eletrônica do peptídeo permitiu modelar 4 resíduos (LYFQ). Já na molécula B, foi possível modelar seis resíduos (ENLYFQ), porém, os resíduos a mais foram estabilizados por interações com a molécula A. Sendo assim, o que melhor representa ligante de um monômero é o que foi observado na molécula A (LYFQ).

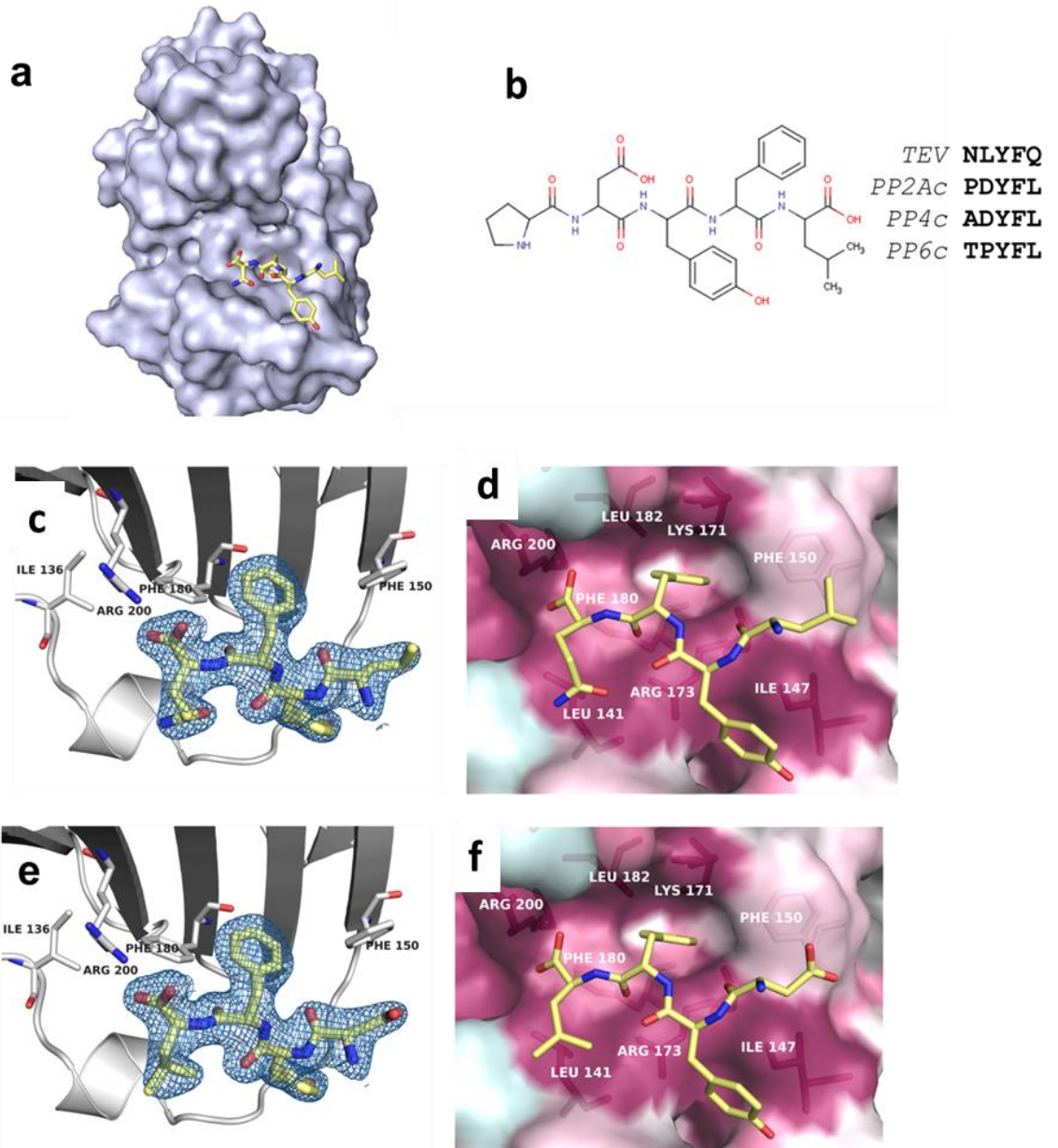


**Figura 3.17:** Organização da TIPRLΔN dentro do cristal. (a) Interface do dímero na região N-terminal corroborada pelo *cross-linking* MS entre os resíduos K29 (b) ponte dissulfeto entre as C87 na porção C-terminal. (c) Organização do cristal e o canal de solvente, as esferas em vermelho representam as pontes dissulfetos entre as C87 e a porção em verde a interface de dimerização.



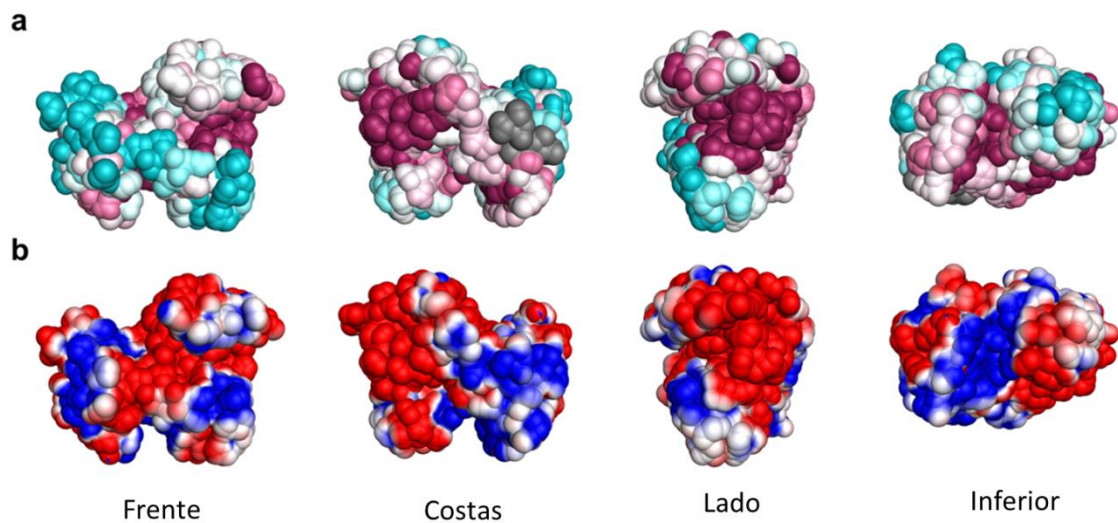
**Figura 3.18:** Estrutura global da proteína TIPRLΔN. (a) Representação em *cartoon* do *fold* da TIPRLΔN destacando seus elementos estruturais. (b) Diagrama topológico do enovelamento da TIPRLΔN gerada com *TopDraw*. (c) e (d) Elementos de estrutura direção N para o C-terminal.

O peptídeo LYFQ do sítio de clivagem da TEV é, estruturalmente, muito semelhante ao C-terminal das proteínas fosfatases PP2Ac, PP4 e PP6 formado pelos resíduos DYFL. Então, modelou-se na densidade de diferença o C-terminal da PP2Ac e se encaixou perfeitamente (Figura 3.19). Ao analisar este modelo, verificou-se que não há espaço para um peptídeo fosforilado na tirosina DpYFL para o encaixe ser bem sucedido. Todavia, há espaço no fundo da cavidade, onde foram observados os resíduos do sítio de clivagem da TEV, para um peptídeo metilado no último resíduo de aminoácido leucina. A cavidade de ligação com o peptídeo é formada por resíduos conservados hidrofóbicos (I147 / F150 e L141 / L182) e resíduos carregados positivamente (K171 / R173 e R184 / R200) (Figura 3.20) e por isso, esses resíduos foram selecionados para os experimentos de mutagênese sítio-dirigida.



**Figura 3.19:** Cavidade da TIPRL $\Delta$ N onde abriga o peptídeo **(a)** Visão geral dos últimos quatro resíduos do peptídeo gerado após a clivagem com a protease de TEV (MGHHHHHHENLYFQ), LYFQ, ligado a TIPRL **(b)** Semelhança entre as sequências do peptídeo do sítio de clivagem da TEV e o C-terminal conservado das fosfatases tipo 2A (PP2Ac, PP4c e PP6c). **(c)** e **(d)** Peptídeo, LYFQ, provindo da clivagem da cauda de histidina modelado no mapa de densidade eletrônica da TIPRL $\Delta$ N e o encaixe perfeito dentro da cavidade. **(e)** e **(f)**. C-terminal da PP2Ac, DYFL, modelado entro do mapa de densidade eletrônica e seu encaixe na cavidade.





**Figura 3.20:** Propriedades da superfície da TIPRL $\Delta$ N. **(a)** Padrão de conservação da sequência TIPRL obtido utilizando o servidor Consurf. A representação da superfície é a partir da cor menos conservada (ciano) para os resíduos mais conservados (magenta). **(b)** Representação da superfície da macromolecular, cálculo da superfície eletrostática feito através do programa PyMOL APBS (*Adaptive Poisson-Boltzmann Solver*)

### 3.3.5 *Cross-linking* acoplado a MS

Na técnica de *cross-linking* acoplado a MS (Paramelle *et. al.*, 2013; Santos *et. al.*, 2011; Fioramonte *et. al.*, 2012) utiliza-se um reagente químico que se liga, preferencialmente, a grupos aminas primárias livres como N-terminal e lisinas, e promove ligações intrapeptídeos entre resíduos de aminoácidos pertencentes ao mesmo peptídeo, bem como ligações interpeptídeos entre resíduos de aminoácidos de peptídeo diferentes. Estas ligações, consequentemente, podem ser classificadas como intraproteína ou interproteína, onde a ligação entre peptídeos pelo *cross-linker* ocorre na mesma proteína/cadeia ou em proteínas/cadeias diferentes, respectivamente. Os espectros são interpretados e obtém-se uma lista de vínculos intra- e interpeptídeos. Esses vínculos de distância podem, por sua vez, ser utilizados para gerar modelos computacionais mais próximos da realidade. Os estudos de *cross-linking* e de troca de hidrogênio por deutério realizados para a proteína TIPRL $\Delta$ N e TIPRL *Full* foram feitos com colaboração do Prof. Dr. Fábio C. Gozzo e sua orientada de doutorado Tatiani Brenelli de Lima do Instituto de Química/Unicamp.

Foram realizadas três análises de *cross-linking*-MS, sendo duas para a proteína TIPRLΔN e uma para a TIPRL *Full*. Como a TIPRLΔN forma dímeros quando a cauda de histidina é clivada e em altas concentrações, mais de um experimento foi planejado, com diferentes construções, para distinguir vínculos interpeptídeos no mesmo monômero de vínculos interpeptídeos da interface de dimerização.

Experimentos com a TIPRL *Full*, que ocorre exclusivamente como monômero, foram feitos para confirmar quais vínculos são do monômero e quais são do dímero, assim pode-se utilizar as informações para propor a interface de interação entre o dímero desta proteína, já que essa informação não se obtém na cristalografia. A Tabela 3.6 traz os vínculos da reação de *cross-linking* obtidos. Os vínculos que estão presentes na TIPRL *Full* são da forma monomérica.

Foram obtidos 17 vínculos interpeptídeos que restringem a distância topológica (que depende do enovelamento local) entre os nitrogênios das lisinas a um máximo em torno de 11 Å. Com respeito à numeração da Tabela 3.6, sete vínculos são exclusivos para o monômero (coluna TF). Outros nove (3-8; 14-16) podem ser de monômero ou dímero, sendo que o último vínculo da Tabela (vínculo de número 17) ocorre entre o mesmo peptídeo, necessariamente, um vínculo na interface do dímero.

Os vínculos concentram-se principalmente em regiões móveis, principalmente no C-terminal, para o qual não é observada densidade eletrônica na estrutura cristalográfica por ser uma região flexível. Um outro ponto é que, segundo a literatura, estima-se que o número de restrições de distância que permite determinar corretamente o *folding* de uma proteína e modelar sua estrutura de forma suficientemente acurada esteja em torno de 1/8 do número total de resíduos de aminoácidos (Roy *et. al.*, 2010; Li *et. al.*, 2004). No caso da TIPRLΔN (260 resíduos), em torno de 30 vínculos. Portanto, o número de vínculos não foi suficiente para modelar a proteína. Esses estudos foram realizados antes de obter a estrutura cristalográfica da proteína TIPRLΔN.

**Tabela 3.6:** Vínculos de *cross-linking*-MS para as amostras de TIPRL. As colunas T1 e T2 correspondem ao primeiro e segundo experimento. TF é o experimento com a TIPRL *Full*. Os resíduos de lisina, em vermelho, sofreram reação de *cross-linking*. Inclui-se, também, a sequência dos peptídeos que incluem o resíduo.

N	Vínculos Interpeptídeos	T1	T2	TF
1	<sup>127</sup> LKVVPTTDHIDTEK <sup>140</sup> - <sup>246</sup> EAVCEK <b>L</b> LIFPER <sup>257</sup>	X		X
2	<sup>201</sup> LYHEADK <b>T</b> YMLR <sup>212</sup> - <sup>246</sup> EAVCEK <b>L</b> LIFPER <sup>257</sup>	X	X	X
3	<sup>141</sup> L <b>K</b> AR <sup>144</sup> - <sup>201</sup> LYHEADK <b>T</b> YMLR <sup>212</sup>		X	
4	<sup>127</sup> LKVVPTTDHIDTEK <sup>140</sup> - <sup>141</sup> L <b>K</b> AR <sup>144</sup>		X	
5	<sup>28</sup> MKSADVEK <sup>35</sup> - <sup>141</sup> L <b>K</b> AR <sup>144</sup>		X	
6	<sup>28</sup> MKSADVEK <sup>35</sup> - <sup>201</sup> LYHEADK <b>T</b> YMLR <sup>212</sup>		X	
7	<sup>141</sup> L <b>K</b> AR <sup>144</sup> - <sup>246</sup> EAVCEK <b>L</b> LIFPER <sup>257</sup>		X	
8	<sup>127</sup> LKVVPTTDHIDTEK <sup>140</sup> - <sup>201</sup> LYHEADK <b>T</b> YMLR <sup>212</sup>		X	
9	<sup>127</sup> LKVVPTTDHIDTEK <sup>140</sup> - <sup>258</sup> IDPNPADSQ <b>K</b> STQVE <sup>272</sup>		X	X
10	<sup>201</sup> LYHEADK <b>T</b> YMLR <sup>212</sup> - <sup>258</sup> IDPNPADSQ <b>K</b> STQVE <sup>272</sup>		X	X
11	<sup>246</sup> EAVCEK <b>L</b> LIFPER <sup>257</sup> - <sup>258</sup> IDPNPADSQ <b>K</b> STQVE <sup>272</sup>		X	X
12	<sup>12</sup> DFCFGPW <b>K</b> LTASK <sup>24</sup> - <sup>258</sup> IDPNPADSQ <b>K</b> STQVE <sup>272</sup>			X
13	<sup>83</sup> L <b>K</b> VACAEWQESR <sup>95</sup> - <sup>218</sup> ES <b>K</b> ISSLMHVPPSLFTEPNEISQYLPIK <sup>245</sup>			X
14	<sup>141</sup> L <b>K</b> AR <sup>144</sup> - <sup>258</sup> IDPNPADSQ <b>K</b> STQVE <sup>272</sup>		X	
15	<sup>25</sup> THIM <b>K</b> SADVEK <sup>35</sup> - <sup>258</sup> IDPNPADSQ <b>K</b> STQVE <sup>272</sup>		X	
16	<sup>118</sup> G <b>T</b> LLGESL <b>K</b> LK <sup>128</sup> - <sup>258</sup> IDPNPADSQ <b>K</b> STQVE <sup>272</sup>		X	
17	<sup>25</sup> THIM <b>K</b> SADVEK <sup>35</sup> - <sup>25</sup> THIM <b>K</b> SADVEK <sup>35</sup>		X	
<u>Vínculos Intrapeptídeos</u>				
18	<sup>96</sup> TEGEHS <b>K</b> EV <b>K</b> IPYDWTYTTDYK <sup>117</sup>	X		X
19	<sup>118</sup> G <b>T</b> LLGESL <b>K</b> L <b>K</b> VVPTTDHIDTEK <sup>140</sup>			X
20	<sup>129</sup> VVPTTDHIDTE <b>K</b> L <b>K</b> AR <sup>144</sup>			X
21	<sup>218</sup> ES <b>K</b> ISSLMHVPPSFTEPNEISQYLPI <b>K</b> EAVCEKLIPER <sup>255</sup>			X

### 3.3.6 Troca de Hidrogênio por Deutério

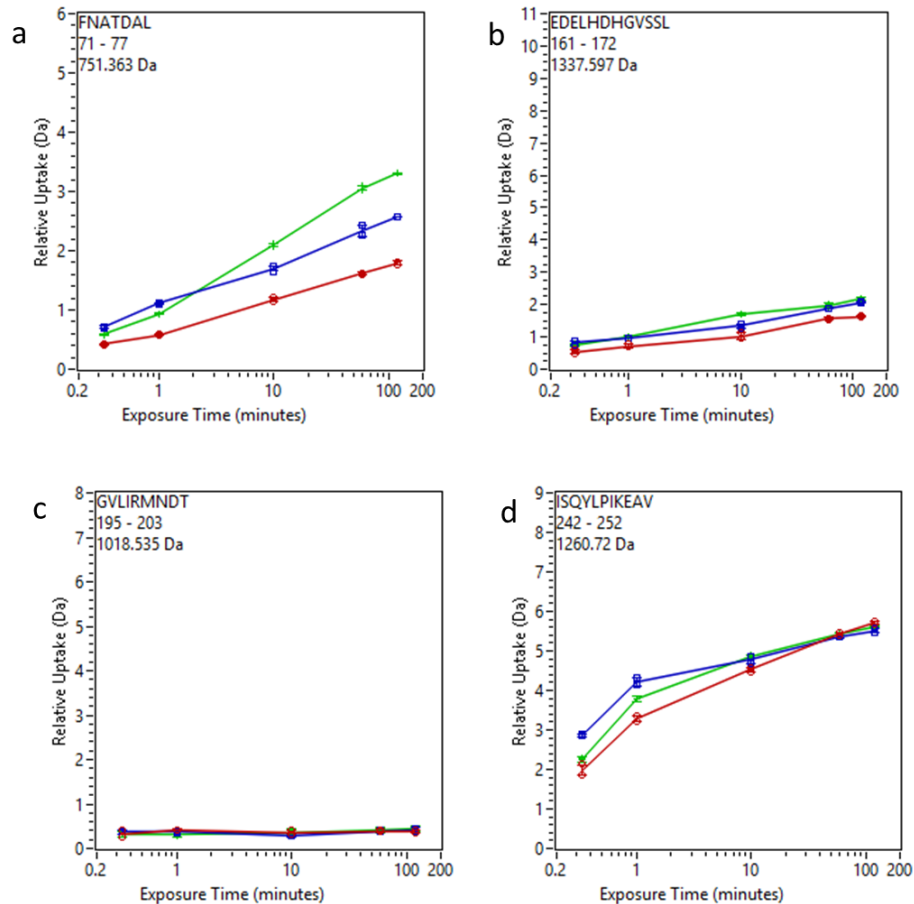
O experimento de troca de hidrogênio por deutério foi realizado com as amostras TIPRL *Full*, TIPRLΔN monômero e dímero para analisar o comportamento da estrutura em solução.

Os resíduos de aminoácidos 160 a 203 que estão localizados no *core* de folha-β trocaram pouquíssimos hidrogênios por deutérios no experimento de troca de hidrogênio por deutério indicando ser uma região rígida. O resíduo de aminoácido

D71, que é extremamente conservado e está localizado em uma  $\alpha$ -hélice exposta ao solvente, realizou uma quantidade considerável de trocas (Figura 3.21). Os resíduos de aminoácidos que formam o extremo C-terminal apresentam muitas trocas de hidrogênio por deutério, indicando uma região altamente flexível e exposta ao solvente. Isso está de acordo com a ausência de densidade eletrônica na porção do extremo C-terminal da estrutura cristalográfica. Além disso, os dados dos experimentos de *cross-linking* acoplado a MS e confirmam essa flexibilidade do C-terminal, já que houve *cross-linkings* do C-terminal com várias porções da proteína.

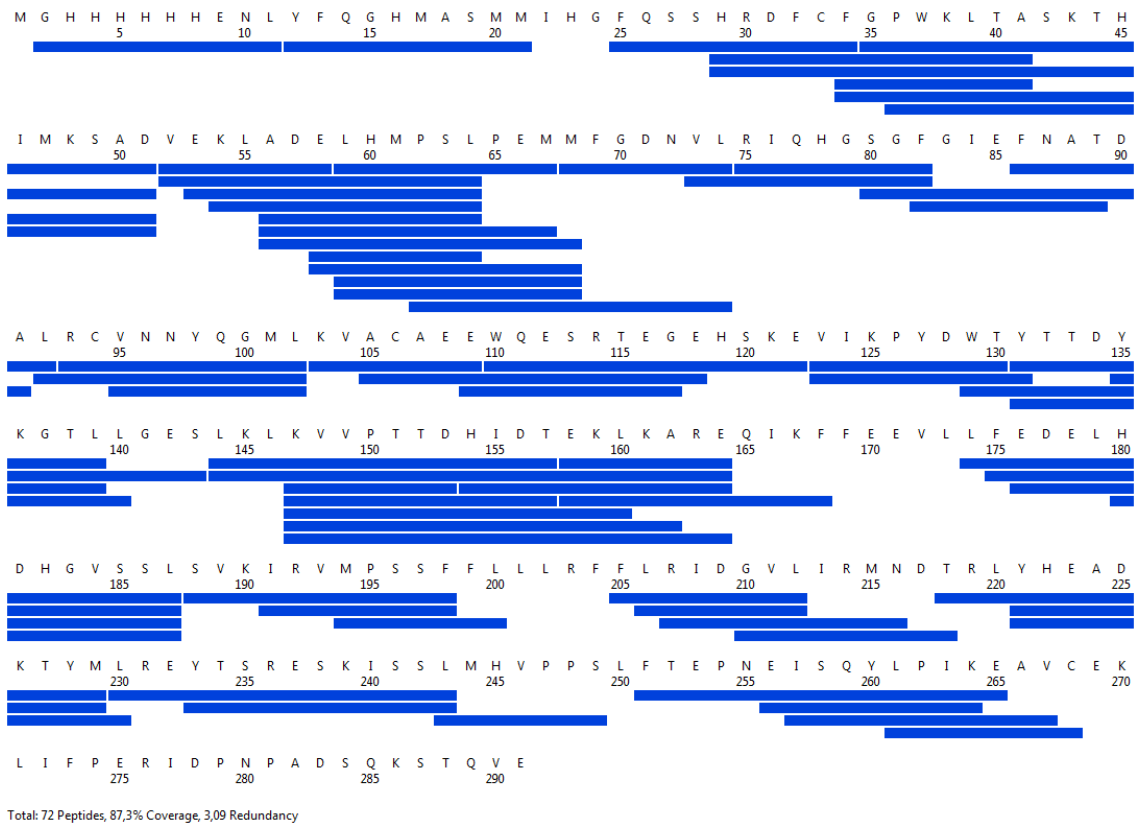
Diante dos resultados expostos na seção 3.3.4, era necessária uma validação experimental sobre a possibilidade de que o peptídeo formado a partir da clivagem com TEV tivesse o mesmo efeito do C-terminal da PP2Ac na TIPRL $\Delta$ N, uma vez que ambos peptídeos são semelhantes. A abordagem ideal para checar essa interação seria obter a estrutura cristalina de um complexo de TIPRL $\Delta$ N com a sequência derivada de PP2Ac em vez do peptídeo, no entanto, não seria viável fazer *soaking* nos cristais com um peptídeo sintético da PP2Ac, pois o sítio já estaria ocupado com o peptídeo. Para obter tal estrutura cristalográfica seria necessário mudar de vetor de expressão e redefinir todas as metodologias aplicadas, pois essa mudança na sequência expressa poderia modificar a proteína. Então, buscaram-se as alternativas viáveis para checar a interação entre a TIPRL e o C-terminal da PP2Ac. Essas abordagens foram a troca de hidrogênio por deutério para validar a ligação da PP2Ac no sítio da TIPRL e experimentos de mutagênese sítio-dirigida associada a GST-*pull-down*.

A análise de troca de hidrogênio-deutério foi proposta para obter informações sobre a flexibilidade e dinâmica da proteína TIPRL *Full* no complexo com os peptídeos sintéticos DYFL e DpYFL, peptídeo fosforilado na tirosina. Estas informações foram obtidas a partir da taxa de incorporação de deutério da cadeia principal, isto é, os átomos de hidrogênio da cadeia principal que estão envolvidos em ligações de hidrogênio e/ou em regiões pouco acessíveis ao solvente apresentam taxas de incorporação de deutério menores quando comparados com os átomos de hidrogênio que estão acessíveis ao solvente e/ou em regiões flexíveis.



**Figura 3.21:** Flexibilidade da estrutura da TIPRL $\Delta$ N. Experimento de HDX onde se observa a flexibilidade das regiões (a)  $\alpha$ -hélice onde está o resíduo de aminoácido D71, houve uma troca considerável de hidrogênios por deutérios. (b) e (c) região conservada que compõem domínio C de folha- $\beta$  formada por uma região conservada de resíduos de aminoácidos 160-203, quase não houve trocas, indicando o quão é rígida essa porção. (d) C-terminal da proteína, essa região não apresentou densidade eletrônica na estrutura indicando sua flexibilidade e confirmada pelo HDX, pois trocou seis átomos de hidrogênios por deutérios ao longo do tempo. As linhas em vermelho são da proteína TIPRL *Full*, azul monômero e em verde dímero da TIPRL $\Delta$ N.

Inicialmente foram realizados ensaios com diferentes concentrações de proteína para otimização da cobertura da sequência de resíduos de aminoácidos. Esta otimização resultou em um mapa de cobertura péptico de 87% dos resíduos de aminoácidos da sequência (Figura 3.22).

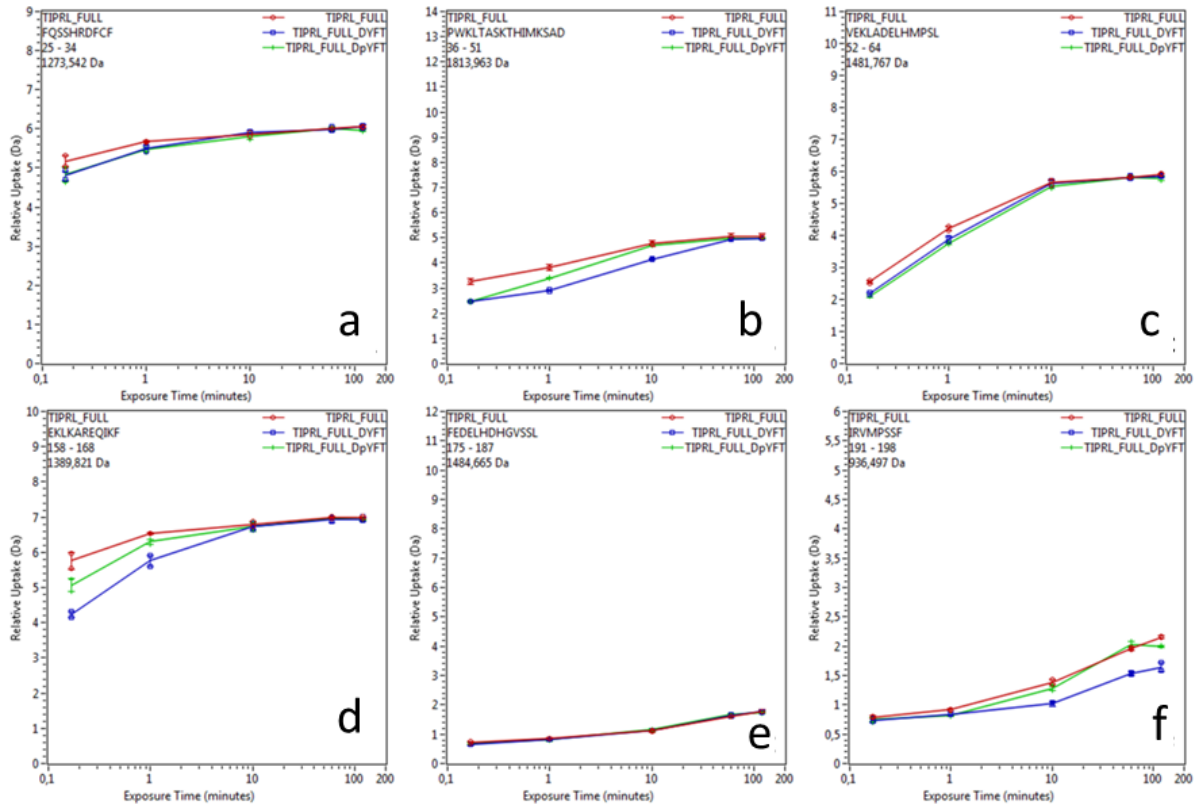


**Figura 3.22:** Mapa de cobertura da sequência de resíduos de aminoácidos da TIPRL *Full* obtidas pela otimização do ensaio de HDX. Os peptídeos identificados após a digestão péptica e análise LC-MS<sup>E</sup> estão demonstrados como faixas em azul obtendo uma cobertura final de 87 %.

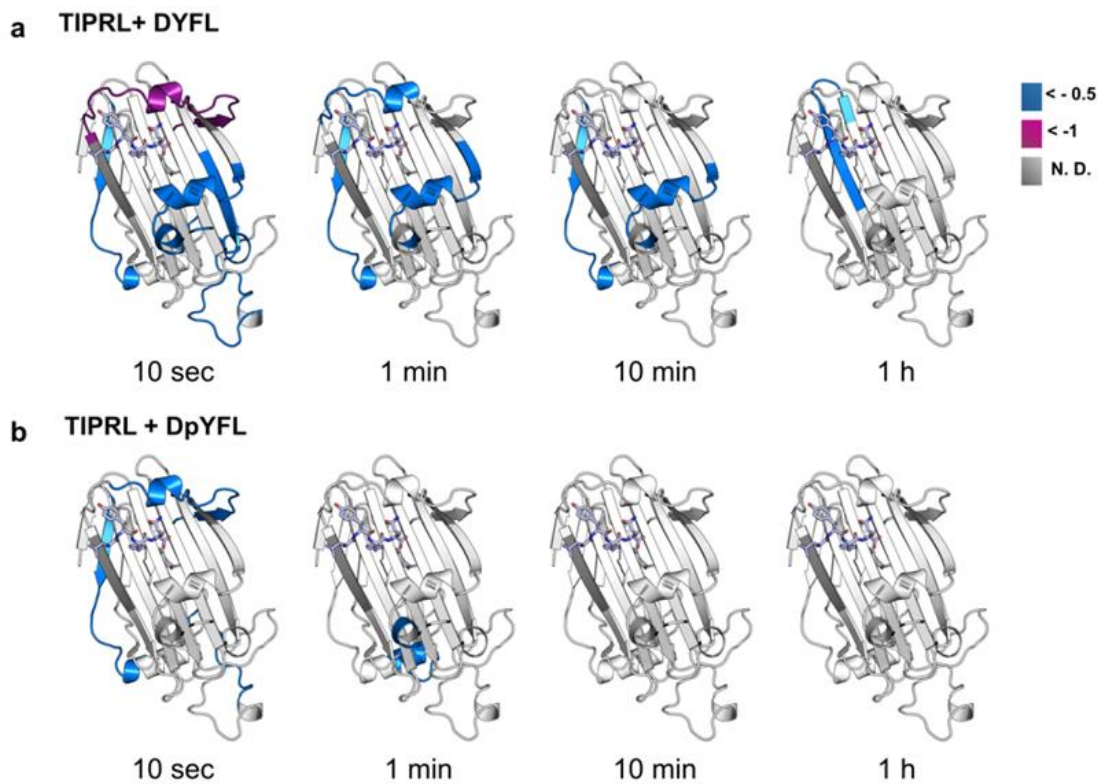
Para determinar a taxa de incorporação de deutério para os diferentes peptídeos foi utilizado o *software DynamX*, onde para cada peptídeo é calculada a massa média a partir do padrão isotópico e, através da massa média dos peptídeos deuterados e do peptídeo no controle, é calculada a taxa de incorporação relativa de deutério.

Para obter informações quanto à dinâmica conformacional da TIPRL *Full* na presença e ausência de ligante, foram comparadas as taxas de deuteração relativas nas diferentes condições. Após esta comparação foi possível observar que a maioria dos peptídeos apresentou estas taxas semelhantes em todas as condições (Figura 3.23 a c e). No entanto, para alguns peptídeos ocorre uma diminuição na taxa de incorporação de deutério para a proteína na presença dos ligantes (Figura 3.23 b d f) indicando, assim, uma diminuição na dinâmica desta região e/ou na acessibilidade ao solvente. Estes peptídeos, com diferenças significativas de incorporação de deutério (considerando uma diferença significativa maior que 0,5

Da), foram correlacionados com a estrutura da TIPRL *Full* na presença de ligante Figura 3.24.



**Figura 3.23:** Representação de alguns gráficos de incorporação de deutério para a TIPRL *Full* na presença de peptídeos ligantes DYFL e DpYFL. Alguns peptídeos apresentaram taxas relativamente semelhantes de deuteração (**a**, **c** e **e**), enquanto outros, demonstraram uma menor taxa de incorporação para a proteína na presença dos ligantes (**b**, **d** e **f**).



**Figura 3.24:** Representação estrutural da análise HDX-MS da TIPRL *Full* com cauda de histidina na presença de peptídeos. **(a)** peptídeo representando o C-terminal da PP2Ac e **(b)** peptídeo fosforilado representando o C-terminal da PP2Ac . Taxa de troca de hidrogênio por deutério Azul: 0 a -0,5 Da, roxo: 0,5 para -1,0 Da, cinza escuro: não determinado. Em ambos os casos em valores negativos.

Analisando a Figura 3.24 é possível observar que a proteína TIPRL *Full* apresenta uma diminuição na taxa de incorporação de deutério mais significativa no complexo com o peptídeo DYFL quando comparada com complexo do mesmo peptídeo fosforilado no resíduo de tirosina (DpYFL). Além disto, verifica-se que a região onde ocorre interação com o peptídeo demonstra uma diminuição na taxa de deuteração, uma vez que, na presença do ligante (DYFL ou DpYFL), os átomos de hidrogênio da cadeia principal desta região estão menos acessíveis ao solvente.



### 3.3.7 Mutagênese sítio-dirigida associada a GST-pulldown

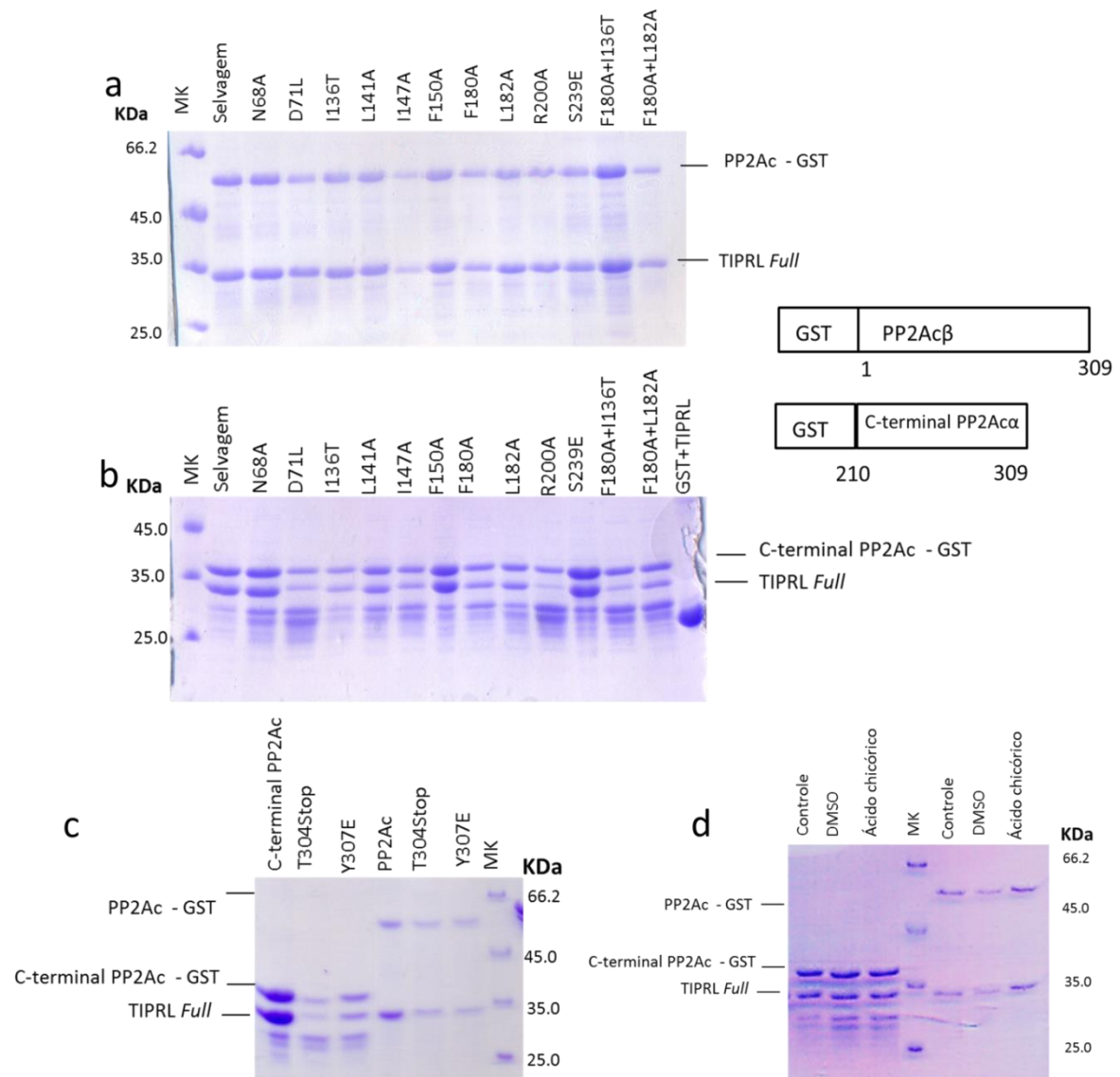
Observando o sítio de interação da TIPRL com peptídeo do sítio de clivagem da TEV e resultados prévios do grupo, foi sugerido que alguns resíduos de aminoácidos fossem mutados na TIPRL *Full*. Alguns estudos já tinham identificado mutações na sequência de TIPRL que prejudicam a sua interação com as subunidades catalíticas de fosfatases tipo 2A (incluindo PP2Ac, PP4c e PP6c). Smetana e Zanchin, 2007, mostraram que os mutantes D71L, I136T, T138S, M196V e D198N alteravam a interação com a fosfatase 2A. Os resíduos I136, T138, M196 e D198 são todos perto do sítio de ligação de peptídeos, o que apoia fortemente a hipótese de que este sítio está envolvido na interação com PP2Ac. Neste mesmo estudo, foi identificada a menor região de interação da PP2Ac, resíduos 210 – 309, com a TIPRL.

Foram selecionados os resíduos de aminoácidos D71 e I136 a partir de resultados prévios do grupo, o resíduo N68, que é conservado e está perto do D71, e os resíduos que formam o sítio L141, I147, F150, F180, R184, L182, R200, além do sítio de possível regulação da TIPRL por fosforilação, S239. A capacidade destas mutações de diminuir a interação com a PP2Ac $\beta$  inteira e com a menor parte que interage com a TIPRL, o C-terminal da PP2Ac (resíduos 210-309), foi avaliada no ensaio de GST-pulldown utilizando proteínas recombinantes coexpressas em *E. coli* (Figura 3.25). Antes de iniciar os ensaios, todos os mutantes foram sequenciados para validar a mutação, e apenas o clone para o mutante R184A não foi validado, sendo assim, foi descartado dos ensaios subsequentes.

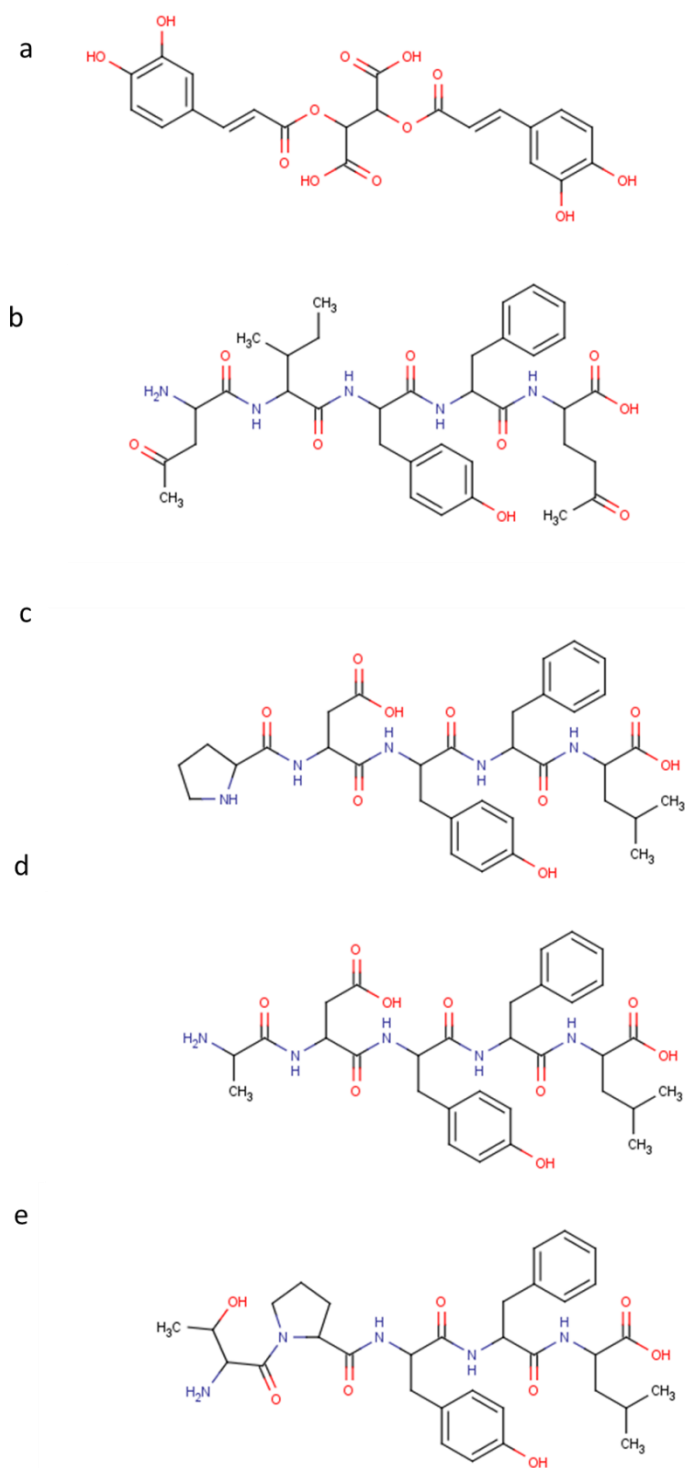
A interação entre a TIPRL e a PP2Ac $\beta$  não foi rompida por qualquer uma das mutações da TIPRL, enquanto a interação entre a TIPRL e a versão truncada da de PP2Ac (resíduos 210-309) foi prejudicada pelos mutantes F180A/I136T e R200A da TIPRL. Este resultado sugere que exista mais de uma superfície interação TIPRL-PP2Ac que inclui o extremo C-terminal de PP2Ac mas não se limita a ele. Provavelmente, mediante eliminação de uma grande parte da sequência de PP2Ac, a interação de TIPRL com o fragmento de PP2Ac restante é estabilizada por contatos com o extremo C-terminal.

Por outro lado, também foram modificados o extremo C-terminal de PP2Ac deletando os últimos cinco resíduos (PP2Ac 304-*Stop*) ou mutando o resíduo Y307 para glutamato (Y307E) que é o fosfomimético, tanto para o C-terminal da PP2Ac quanto para a PP2Ac inteira e avaliou-se a interação com a TIPRL *Full* selvagem. Estas mutações não tiveram nenhum efeito significativo sobre a interação.

Yoon e colaboradores, 2015, (Yoon *et. al.*, 2015) observaram que a interação do complexo entre TIPRL e MKK7 (quinase presente em várias vias de sinalização celular) sofre influências na presença de extratos de plantas (*Taraxacum officinale* – dente de leão) que continham ácido chicórico, o qual é, estruturalmente, semelhante à porção C-terminal da PP2Ac que interage com o bolsão hidrofóbico da TIPRL (Figura 3.26). A partir dessa informação, realizou-se um GST-*pull-down* entre a PP2Ac e a TIPRL *Full* na presença do ácido chicórico. Porém, nas condições testadas o ácido chicórico não afetou a interação entre as proteínas.



**Figura 3.25:** Mutagênese sítio-dirigida associada a GST-pulldown. **(a)** Interação entre a TIPRL e PP2Ac $\beta$ , **(b)** TIPRL e C-terminal PP2Ac. Possível observar redução na interação com os mutantes R200 e F180/136. Em destaque ao lado, os mutantes da TIPRL Full que afetaram a interação com o C-terminal da PP2Ac comparado com o selvagem. **(c)** PP2Ac com mutações sítio-dirigidas e TIPRL Full selvagem, não houve mudanças na interação. **(d)** GST-Pulldown na presença do ácido chicórico, não houve mudanças na interação entre as proteínas.



**Figura 3.26:** Comparação estrutural entre o ácido chicórico e os peptídeos que interagem com a TIPRL. (a) ácido chicórico (b) peptídeo do sítio de clivagem da TEV (c) C-terminal da PP2Ac (d) C-terminal da PP4c (e) C-terminal da PP6c.

# **CAPÍTULO 4**

**Manuscrito de artigo científico descrevendo os resultados para a proteína TIPRL**

## 4. Manuscrito TIPRL

### Crystal structure of the human Tip41 orthologue, TIPRL, reveals a novel fold and a binding site for the C-terminus of protein phosphatase 2A

Valéria Scorsato<sup>1,2,3</sup>, Tatiani Lima<sup>1</sup>, Germanna L. Righetto<sup>3</sup>, Nilson I. T. Zanchin<sup>4</sup>, José Brandão-Neto<sup>5</sup>, James Sandy<sup>5</sup>, Humberto D’Muniz Pereira<sup>6</sup>, Fabio C. Gozzo<sup>1</sup>, Juliana H. C. Smetana<sup>3</sup> and Ricardo Aparicio<sup>1</sup>

1. Institute of Chemistry, University of Campinas, Campinas, São Paulo 13083-970, Brazil.

2. Institute of Biology, University of Campinas, Campinas, São Paulo 13083-970, Brazil

3. Brazilian National Laboratory for Biosciences, Center for Research in Energy and Materials, Campinas, São Paulo 13084-971, Brazil.

4. Oswaldo Cruz Institute (FIOCRUZ), Curitiba, Paraná, Brazil,

1. Diamond Light Source, UK

2. Physics Institute, University of São Carlos, Brazil,

Correspondence and requests for materials should be addressed to J.H.C.S (e-mail [juliana.smetana@lnbio.cnpem.br](mailto:juliana.smetana@lnbio.cnpem.br)) or R.A. (e-mail [aparicio@iqm.unicamp.br](mailto:aparicio@iqm.unicamp.br)).

### Summary

TIPRL is a regulatory protein which inhibits the catalytic subunits of Type 2A phosphatases. Several cellular contexts have been proposed for TIPRL, such as regulation of mTOR signaling, inhibition of apoptosis and biogenesis and recycling of PP2A, however, the underlying molecular mechanism is still poorly understood. We have solved the crystal structure of human TIPRL at 2.15 Å resolution. The structure is a novel fold organized around a central core of planar, antiparallel beta-sheet, showing an N-terminal  $\alpha/\beta$  “lid” domain at one of its surfaces and a conserved cleft at the opposite surface. Inside this cleft, we found a peptide derived from TEV-mediated cleavage of the affinity tag. We show by mutagenesis, pulldown and HDX/MS that this peptide is a decoy for the conserved C-terminal tail of PP2A, an important region of the phosphatase which undergoes post-translational modifications and regulates holoenzyme assembly. TIPRL preferentially binds the unmodified version of the PP2A-tail mimetic peptide DYFL compared to its tyrosine-phosphorylated version. Our findings provide a structural framework for the function of TIPRL in PP2A inhibition and

place it in the important category of cellular proteins that recognize the C-terminal tail of PP2A, including PTPA, PME-1 and LCMT-1.

## Introduction

Reversible phosphorylation orchestrates most signaling pathways. Type 2A phosphatases, comprising PP2A, PP4 and PP6, are serine/threonine phosphatases of the PPP subfamily which are structurally and evolutionarily related to PP1, while sharing particular distinctive characteristics such as sensitivity to inhibition by okadaic acid in the nanomolar range and the presence of a conserved C-terminal motif (PDYFL) which is subject to post-translational regulation in PP2Ac<sup>1,2</sup> and possibly also in PP4c and PP6c<sup>3</sup>. PP2A is a major regulator of several cellular processes and an important tumor suppressor. Its functional unit is a heterotrimeric holoenzyme composed of a catalytic subunit (PP2Ac), a scaffold subunit (PR65/A) and a regulatory subunit which might belong to four structurally diverse families (B, B', B'' and striatins). The selectivity is mediated by post-translational modifications of the C-terminal tail, which can be phosphorylated on the tyrosine residue or methylated on the extreme C-terminal leucine<sup>4,5</sup>.

The biogenesis of PP2A, defined as the process undergone by the phosphatase catalytic subunit from translation of the polypeptide to its incorporation into functional holoenzymes, is regulated by a series of steps which include stabilization of a latent, inactive form by  $\alpha 4$ <sup>6</sup>, ATP-dependent transfer of catalytic metal ions by PTPA<sup>7</sup> and subsequent methylation of the C-terminal leucine (Leu309) by LCMT-1 (Leucine Carboxy Methyl Transferase)<sup>8</sup>. In a recent review, Sents et al. proposed that five proteins are responsible for the biogenesis of PP2A: LCMT1, PME-1, PTPA,  $\alpha 4$  and potentially TIPRL<sup>9</sup>.

TIPRL (TIP41, TOR signaling pathway regulator-like) is a type 2A phosphatase regulatory protein which is still poorly understood, in part due to the lack of any structural information which might guide further studies of its structure-function relationship. In our previous studies, we performed a yeast two-hybrid screen using human TIPRL as bait and found that it binds the catalytic subunits of PP2A, PP4 and



PP6 and inhibits the activity of PP2A in vitro<sup>10</sup>. A recent study showed that TIPRL is overexpressed in hepatocellular carcinoma and its down-regulation synergizes with TRAIL to induce apoptotic death of tumor cells<sup>11</sup>. TIPRL has been considered a part of the mTOR signaling mostly due to the function of its yeast counterpart, Tip41<sup>12</sup>, but experimental evidence was lacking until recently. However, contrary to the yeast Tip41, which inhibits TOR signaling by sequestering the  $\alpha 4$  orthologue Tap42, human TIPRL positively regulates mTORC1 signaling in response to amino acids<sup>13</sup>.

During the past decade, several studies elucidated the mechanisms of PP2A regulation by solving the crystal structures of its regulatory proteins, either isolated<sup>14-18</sup> or in complex with PP2Ac or the PP2A core enzyme<sup>6-8,19,20</sup>. However, TIPRL was still eluding most efforts and remained largely uncharacterized. Here, we report the crystal structure of human TIPRL bound to a TEV site-derived peptide which unexpectedly functioned as a decoy for the conserved C-terminus of PP2Ac. Our findings provide a structural framework to interpret the inhibition of PP2A by TIPRL.

## Results

### Overall structure of TIPRL

After extensive crystallization attempts using full length TIPRL without success, we decided to use short N-terminal or C-terminal deletions to eliminate potential flexible regions. The C-terminal deletion was deleterious to the integrity of the protein, targeting its expression to inclusion bodies (not shown), but the N-terminal deletion didn't affect the yield of soluble protein in *E. coli*. This construct, named TIPRL $\Delta$ N, was used in the subsequent crystallization experiments. The high solubility of the protein prevented its crystallization at a lower concentration range of 5 – 10 mg ml<sup>-1</sup>, and gradual increases in the concentration of the sample led us to find the nucleation zone around 70 – 100 mg ml<sup>-1</sup>, which produced bipyramidal crystals belonging to *P6<sub>5</sub>* space group with two TIPRL molecules in the asymmetric unit and a solvent content of 76.2 %. These crystals diffracted up to 2.15 Å. We solved the crystal structure by SAD using selenomethionine labelled crystals. Data collection and refinement statistics are shown in Table 1.

Prior to solving the crystal structure of TIPRL, we attempted a computational approach combined with crosslinking analysis to elucidate its structure. We used several structure prediction servers, such as I-Tasser<sup>21</sup>, Robetta, Quick Phyre and Quark, all of which failed to provide a consistent model that could satisfy all the distance restraints. After the first crystals were obtained, initial attempts to solve the structure using molecular replacement for phasing also failed. These observations were suggestive of a novel fold for TIPRL, however unlikely due to the growing scarcity of new folds deposited on PDB. Indeed, the solved structure obtained through selenomethionine experimental phasing was not reminiscent of any known protein fold. Structural similarity searches on DALI, Top Search and CATH did not find significant hits. We therefore assume that the TIPRL/Tip41 family displays a fold that was not previously described.

The structure consists of a planar, antiparallel beta-sheet core with strikingly long strands of up to 12 residues, which is covered by a “lid” domain where a smaller beta-sheet is present along with two alpha-helices and a considerable amount of residues that do not adopt either alpha-helix or beta-strand conformation, yet are ordered enough to appear in the electron density map (Figure 1). The lid domain corresponds to the N-terminus, while the beta-sheet core corresponds to the C-terminus and consists mostly of conserved residues. Both domains are connected through a long loop (residues 121 – 148) containing a short segment in alpha-helical conformation, which embraces the beta sheet core and covers part of its bottom side forming a cleft along with a small, helical subdomain located opposite to it, predominantly formed by  $3_{10}$  helices. For the purpose of further discussion, we define this region as the bottom cleft, the N-terminal region as the upper side, and residues 121 – 148 as the connector loop. The C-terminal beta-sheet domain shows the highest evolutionary conservation, highlighting its structural importance (Figures 1e and 2a). The electrostatic potential shows a predominantly negative surface while the bottom cleft concentrates positively charged residues which are also highly conserved (Figure 2b).

TIPRL $\Delta$ N forms a dimer in solution, while TIPRL is likely a monomer in the cell

The asymmetric unit of the crystal contained two molecules, MolA and MolB, which were refined independently, but did not show any major differences upon completion of the refinement cycles showing an RMSD value of 0.3 (Figure 3). The crystal packing was stabilized by two interfaces, one of them involving the N-termini of both monomers and another one between the unstructured regions of the N-terminal domains of both monomers. A continuous electron density between the side chains of Cys87 on both monomers suggested the presence of a disulfide bond (Figure 3e), which might partially explain the stability of these unusually solvated crystals. Interestingly, a dimer was observed in solution during the purification process, albeit under unusual conditions which do not support the existence of a physiological TIPRL dimer in the cellular environment.

We noticed an unusual behaviour of TIPRL $\Delta$ N upon proteolytic removal of the N-terminal histidine tag by TEV, which correlated with its ability to form crystals. The sample subjected to crystallization screens was previously cleaved with TEV protease to remove its hexahistidine tag, however, the cleavage was usually incomplete resulting in a double band on SDS-PAGE analysis (Figure 3a). Despite its apparent heterogeneity, this preparation consistently crystallized, while several other preparations that were not subjected to TEV cleavage failed to do so. By contrast, when full length TIPRL was cleaved by TEV under the same conditions, the cleavage efficiency was close to 100%, but this construct did not yield any crystals.

Additionally, TIPRL $\Delta$ N, but not full length TIPRL, was able to dimerize after removal of the tag (Figure 3b). Analysis of size exclusion profiles under different conditions allowed us to conclude that dimerization occurs only if these three conditions are satisfied: 1- deletion of the N-terminal residues, 2- cleavage of the histidine tag, and 3- high concentration. We were able to get crystals from highly purified preparations of cleaved, dimeric protein as well as mixed fractions containing both cleaved and uncleaved protein.

Prior to solving the crystal structure, we suspected that the N-terminal region of TIPRL was involved in the dimer interface because the dimerization affects the cleavage of the N-terminally placed histidine tag. Crosslinking-MS analysis of the dimer identified an interchain crosslink in a peptide from the N-terminal region, <sup>25</sup>THIMKSADVEK<sub>35</sub>, which was later mapped onto the crystal structure within a 18.4 Å distance between Lys29 and its counterpart (Figure 3c). The  $\beta$ -strand in the extreme N-terminus of TIPRL $\Delta$ N is critical to form the dimer interface in the asymmetric unit and its amino acid residues originate from a plasmid-derived sequence instead of TIPRL itself (GHMAS), which helps explain the fact that the full length TIPRL construct does not dimerize. These results also exclude the possibility that the dimer observed during the purification process might correspond to the disulfide-stabilized interface instead of the N-terminal interface.

The preformed dimer appears to be a necessary requisite for nucleation instead of an artifact produced by the crystallization conditions. However, due to the conditions described above that must be fulfilled to allow its appearance, we do not believe that this dimer is a physiologically relevant oligomeric form of TIPRL. Despite extensive evidence for the dimeric form of recombinant TIPRL $\Delta$ N, we conclude that the most likely oligomeric state of TIPRL in the cell is a monomer.

### Flexibility of the C-terminal residues

The electron density of the last 12 C-terminal residues is missing from the final map. This region is relatively short and not conserved, but it harbours two experimentally validated phosphorylation sites. Ser265 was detected along with Ser239 as a phosphorylation site for ATM/ATR kinases<sup>22</sup> and Ser268 is present in PHOSIDA as an observed phosphorylated site from mouse brain samples. Therefore, we sought to explore the possible conformation of this particular region using crosslinking-MS data.

The tryptic peptide corresponding to the extreme C-terminus of TIPRL, <sup>258</sup>IDPNPADSQKSTQVE<sub>271</sub>, was ubiquitous in the crosslinking-MS analyses, which helped us to understand its flexible behavior. This peptide was crosslinked to peptides <sup>12</sup>DFCFGPWKLTASK<sub>24</sub>, <sup>25</sup>THIMKSADVEK<sub>35</sub>, <sup>118</sup>GTLLGESLKLK<sub>128</sub>, <sup>127</sup>LKVVPTTDHIDTEK<sub>140</sub>, <sup>141</sup>LKAR<sub>144</sub>, <sup>201</sup>LYHEADKTYMLR<sub>212</sub> and <sup>246</sup>EAVCEKLIFPER<sub>257</sub>, which span the primary sequence from N- to C-terminus. Upon analysis of the three dimensional structure of TIPRL, we observed that although distant in the primary structure, most of these residues are located within 20Å of the C-terminal extremity (Supplementary Figure 1). The exceptions are Lys29 and Lys142, which are located in opposite sides on the TIPRL surface (the N-terminal lid domain and the connector loop, respectively). This suggests that the C-terminus adopts at least two possible conformations, being either on top of the N-terminal domain or pointing downwards to the connector loop.

## Identification of a binding site for the C-terminus of PP2Ac

During the last refinement cycles, we observed a region of positive electron density inside the bottom cleft which could not be accounted for by the modeled TIPRL sequence. After failed attempts to model it as an alternative conformation of the connector loop, we eventually noticed that the extra electron density could be correctly modeled as the four C-terminal residues from the TEV cleavage site, LYFQ, on MolA (Figure 4). The same peptide was bound to the symmetric binding site on MolB and stabilized by further interactions with MolA, which allowed for the modelling of the entire TEV site, ENLYFQ. This finding implied that the peptide generated after cleavage with the TEV protease (MGHHHHHHENLYFQ) shows a high affinity for TIPRL, which allowed it to stay tightly bound to the protein during the size exclusion and buffer exchange steps instead of being lost by diffusion. Most of the N-terminal region of this peptide is exposed to the solvent-filled channel of the crystal in a random conformation and therefore does not appear in the electron density, while the last four to six residues including the extreme C-terminus are bound to TIPRL.

The binding site for the peptide shows alternating patches of hydrophobic (I147/F150 and L141/L182) and positively charged residues (K171/R173 and R184/R200). Most of the TIPRL-peptide interface is stabilized by hydrophobic interactions. One major exception is Arg200, which engages in ionic interactions with the carboxy moiety of the extreme C-terminus. These residues are highly conserved, suggesting a functional binding site rather than a spurious artifact (Figure 4c). We therefore sought to interpret this finding in the context of the biological function of TIPRL. Upon careful examination of the TEV recognition site, we found that its sequence resembles the conserved C-terminus of PP2Ac, which is also shared by the related phosphatases PP4 and PP6 (Figure 4d). Actually, at the resolution of 2.15 Å the LYFQ peptide could hardly be distinguished from the <sub>306</sub>DYFL<sub>309</sub> sequence of the PP2Ac C-terminus, which allowed us to model the PP2Ac sequence using the electron density from the TEV site peptide (Figure 4e and 4f). By analyzing this model, we noticed that phosphorylation

on Tyr307 would result in steric hindrance, while methylation of the extreme C-terminus would be easily accommodated. This observation raised an intriguing possibility for the regulation of PP2Ac binding to TIPRL.

As Nakashima et al. reported recently<sup>13</sup>, TIPRL preferentially binds methylated PP2Ac and a phosphomimetic mutation of residue Tyr307 to glutamate on PP2Ac prevented its interaction with TIPRL, which the authors interpreted as evidence for regulation based on the methylation status of the C-terminus. Based on our structural analysis, it's more likely that this mutation directly affects the interaction rather than indirectly interfering with the methylation of the C-terminus.

#### Mutational analysis of the TIPRL-PP2Ac interaction and HDX

The possibility that the TEV site peptide might play the role of the PP2Ac C-terminus required strong experimental validation. The ideal approach would be to obtain the crystal structure of a complex of TIPRL with the PP2A-derived sequence instead of the TEV site peptide, however, it would not be feasible to soak the crystals with a synthetic peptide or even incubate this peptide with the purified protein because the binding sites are already occupied with the TEV site peptide, probably during the whole purification procedure subsequent to the cleavage with TEV protease. To obtain such structure we would need to design a completely different purification strategy and perform new crystallization screenings which might not be successful. Instead, we used a combined approach of mutational analyses and hydrogen-deuterium exchange to validate the PP2Ac binding site on TIPRL and to explore the potential role played by the post-translational modifications in the regulation of this interaction.

Our previous studies had identified mutations in the sequence of TIPRL that impaired its interaction with the catalytic subunits of type 2A phosphatases (including PP2Ac, PP4c and PP6c) using a random, unbiased reverse two-hybrid screen<sup>10</sup>. The mutants D71L, I136T, T138S, M196V and D198N were deficient in phosphatase binding as measured by a yeast two-hybrid assay. In the same study, we also showed that residues

210 – 309 of PP2Ac were sufficient to bind TIPRL both in the yeast two-hybrid and in GST pulldown assays. Interestingly, residues I136, T138, M196 and D198 are all close to the peptide binding site, which strongly supports our hypothesis that this site is involved in the interaction with PP2Ac. D71 is located in an opposite region, as discussed above.

We selected two mutants from our previous study (D71 and I136) and performed designed mutations in other residues based on their sequence conservation and presence in the putative PP2Ac binding site. Residue N68, which is conserved and close to D71, was also mutated. The ability of these mutations to disrupt the interaction with either full-length or N-terminally truncated PP2Ac (residues 210 – 309) was assessed in a GST-pulldown assay using recombinant proteins co-expressed in *E. coli* (Figure 5). Intriguingly, the interaction with full length PP2Ac could not be disrupted by any of the TIPRL mutations, while mutants F180/I136, and R200 showed remarkably impaired binding to the truncated version of PP2Ac (residues 210 – 309). This finding suggests an extensive TIPRL-PP2Ac interaction surface which includes the extreme C-terminus of PP2Ac but is not restricted to it. Probably, upon deletion of a major part of the PP2Ac sequence, the interaction of TIPRL with the remaining PP2Ac fragment is stabilized by contacts with the extreme C-terminus.

Conversely, we modified the extreme C-terminus of PP2Ac by deleting the last five residues (PP2Ac 304-Stop) or changing Tyr307 to phosphomimetic Glutamate (Y307E), both on full length and C-terminal (210 – 309) constructs, and assessed the interaction with wild type TIPRL by GST pulldown. These mutations had no appreciable effect on the interaction. (Figure 5d).

Additional insight was obtained from HDX/MS experiments in which purified TIPRL was incubated with the synthetic DYFL peptide, either phosphorylated on the tyrosine residue or unmodified. These four residues which mimic the extreme C-terminus of PP2Ac were sufficient to cause appreciable reduction of backbone amide deuterium intake on TIPRL compared to control sample without ligand, indicating a gain of



secondary structure and reduced solvent exposure of specific regions of TIPRL upon peptide binding. The time course analysis of TIPRL - DYFL peptide after deuterium removal showed that within 10 seconds, the connector loop showed the lowest backbone amide deuterium intake per residue. The same experiment performed with phosphorylated peptide (DpYFL) showed similar effects but with reduced magnitude and shorter duration, indicating that phosphorylation interferes negatively with the interaction. Taken together, the mutagenesis/pulldown and HDX experiments suggest that the extreme C-terminus of PP2Ac is sufficient, but not necessary to bind TIPRL, and support the hypothesis that the binding site for TEV-derived peptide in the crystal structure actually indicates where the PP2A tail would bind.

## Discussion

In the present work we described an intriguing novel fold for TIPRL which is likely the first of a family including all TIPRL/Tip41 related proteins. When we analyzed this new structure, the most relevant of our findings was the identification of a PP2Ac binding site on the surface of TIPRL and characterization of its binding specificity to the conserved C-terminus of PP2Ac. The residues on this binding surface are extremely conserved, as expected for this ancestral function which is shared with the yeast orthologue Tip41.

Previous approaches to map the PP2Ac binding site on TIPRL relied on random mutagenesis<sup>10</sup> or deletion mapping using a naturally occurring shorter isoform. The short isoform of TIPRL is 178 residues long, comprising residues 1 to 172 of the full length protein, with an additional C-terminal extension of 6 amino acid residues (PGGGHL). McConnell et al<sup>23</sup> reported that this isoform does not interact with PP2Ac and therefore concluded that the binding site is located in the C-terminal region of TIPRL (residues 173 – 272), which is consistent with our own findings reported here. However, we noticed that the short isoform might not be correctly folded, as its C-terminus corresponds to the middle region of the beta-sheet core of the C-terminal domain. Our attempts to express and purify this isoform from *E. coli* resulted in insoluble protein under all conditions tested, suggesting that truncation of the TIPRL sequence at that specific region deeply disturbs its structure (not shown). Not surprisingly, this isoform has never been detected as an endogenous protein.

The insight into the PP2A binding site came from the unexpected presence of a TEV site peptide bound to TIPRL. If it is true that a peptide derived from a widely used sequence such as the TEV site can mimic the C-terminus of PP2Ac, then we would expect to see this happening in other structures of PP2Ac-binding proteins. We searched the literature for similar cases and found evidence for this claim by comparing two PDB entries, 4NY3 and 2G62. In the crystal structure of PTPA

reported by Magnusdottir et al.<sup>24</sup>, seven amino acid residues from a TEV-linker were found inside a conserved hydrophobic pocket, which was interpreted as evidence for a protein binding function, however, the authors did not mention the similarity with the PP2A C-terminus. We noticed that the same region of PTPA engages in interaction with the extreme C-terminus of PP2Ac in the crystal structure recently reported by Löw et al.<sup>25</sup>. It is therefore reasonable to argue that the TEV site reproduces the binding specificity of the PP2A C-terminus under different circumstances and even adopting different conformations.

We validated the potential peptide-binding cleft on TIPRL by mutating each one of its highly conserved residues and assessing the interaction with PP2Ac by GST-pulldown assays, as well as by measuring direct binding of a synthetic PP2Ac derived peptide to TIPRL by HDX/MS. While no single mutation on TIPRL was able to reduce the interaction with full length PP2Ac, mutation of residues F180/I136 and R200 showed major effects on its binding to a truncated version of PP2Ac lacking ~200 residues from the N-terminus. The HDX analysis showed that the structural changes induced by peptide binding occur mostly on the helical regions on both sites of the proposed peptide binding cleft, but were not restricted to it. Intriguingly, a major conformational change was observed in residues 74–83, suggesting a long range communication between the bottom cleft and helix  $\alpha 2$  which is mediated by the connector loop. To a lesser extent, similar results were observed with the phosphorylated peptide. Helix  $\alpha 2$  contains the conserved amino acid residue Asp71 which was shown by our previous studies<sup>10</sup> and other's<sup>13</sup> to interfere with PP2Ac binding. Due to its position on the opposite side of the peptide binding cleft, helix  $\alpha 2$  is unlikely to engage in direct interaction with any part of PP2Ac. Instead, it might function as a regulatory switch, either by engaging in interaction with other proteins, or responding to phosphorylation of nearby residues. The mechanism of TIPRL function might rely on this dynamic communication between helix  $\alpha 2$  and the peptide binding cleft, mediated by the unique conformation of the connector loop. However, what triggers this communication in vivo remains to be elucidated.

The N-terminus of the bound peptide points to a highly conserved region at the back side of TIPRL, which suggests a continuity of the PP2A-binding surface from the flexible C-terminus binding site at the bottom cleft of TIPRL to the compact globular domain of the phosphatase which potentially binds to its back side. Interestingly, in the crystal structure this surface is hidden in the dimer interface. This might further indicate that it has a potential for protein binding. The previously reported structure of Tap42 was also a dimer<sup>26</sup> and further studies with the human orthologue  $\alpha 4$  showed that the homodimerization interface actually indicated the PP2A binding site<sup>18</sup>.

The structures of PP2Ac complexed with LCMT1<sup>8</sup> and PME-1<sup>19</sup>, as well as combined structural information from separate PTPA-containing structures<sup>24,25</sup>, show that these proteins bind specifically to the C-terminal tail while engaging in extensive interactions with the phosphatase active site. The fact that TIPRL binds to the PP2A tail indicates that it probably occludes the active site of PP2Ac, which provides a structural basis for its known function as a PP2A inhibitor. Considering that specific binding to the conserved PP2Ac tail is a signature of proteins that participate in the biogenesis/ recycling of the phosphatase, our findings provide a structural framework to further explore the potential role of TIPRL in the assembly and recycling of PP2A holoenzyme complexes.

## Methods

Plasmid construction and mutagenesis. The human TIPRL cDNA (access number NP\_690866.1) was amplified from a fetal brain cDNA library (Clontech) and cloned into pET-TEV, a modified pET28a vector harboring the TEV site, in frame with an N-terminal hexa-histidine tag. The TIPRL $\Delta$ N construct was obtained by amplifying the cDNA from pET-TEV-TIPRL using a forward primer which promoted a 15-residue N-terminal deletion and subcloning the resulting PCR product into pET-TEV. The mutants were obtained by thermal cycling with *Pfu* (Thermo Scientific) and digestion of parental DNA with *DpnI* (Thermo Scientific), followed by transformation into

DH5 $\alpha$  and sequencing. The sequences of primers used for mutagenesis are shown in **Supplementary Table 1**.

Purification and crystallization of TIPRL $\Delta$ N. His-tagged TIPRL $\Delta$ N was expressed in *E. coli* BL21(DE3) for 5 hours at 25°C. Expression was induced with 14 mM lactose. The initial purification step was affinity chromatography with Ni Sepharose 6 Fast Flow (GE Healthcare). The affinity purified protein was cleaved with TEV protease for 16 hours at 4°C and subjected to a final polishing step of size exclusion chromatography on a Superdex 200 26/60 column (GE Healthcare) in 10 mM pH 7.5 Tris buffer, 150 mM NaCl, 10 mM  $\beta$ -mercaptoethanol. The peak fractions were pooled, the buffer was changed to 20 mM pH 7.5 Tris; 25 mM NaCl, 10 mM  $\beta$ -mercaptoethanol. The peak fractions were pooled, the buffer was changed to Tris 20 mM pH 7.5; NaCl 25 mM,  $\beta$ -mercaptoethanol 10 mM using a centrifugal filter device (Millipore, MWCO 10.000) and the sample was concentrated to 70–100 mg ml<sup>-1</sup> using the same device. Selenomethionine labelled protein was expressed in minimal medium M9 supplemented with selenomethionine using 0.5 mM IPTG as inducer, and purified following essentially the same procedure as the native protein. The initial crystallization attempts were performed in the high throughput crystallization facility of the Brazilian National Biosciences Laboratory, RoboLab. Crystals were grown by vapor diffusion at 18 °C. Diffraction quality crystals were grown after 10 days in two conditions: 100 mM di-sodium di-potassium phosphate pH 6 to 6.5; 2 to 3 M NaCl, and 100 mM SPG buffer (Succinic Acid, Sodium Dihydrogen Phosphate and Glycine) pH 5.5 – 6.7.

Structure solution and refinement. Data collection was performed on beamline I02 at the Diamond Light Source (UK). Initial data processing was done using the FAST\_DP pipeline and was followed up by automated phasing using FAST\_EP. Automated model building was carried out with ARP/wARP and resulted in an initial model where 96% of amino acid residues were correctly modeled with  $R_{\text{free}} = 0.22$  at 2.66 Å

resolution. Strong anomalous signal was found during data reduction, which was subsequently confirmed by ShelxC. Following the initial analysis, different ShelxD jobs were run for each of the possible space groups:  $P6$ ,  $P6_1$ ,  $P6_2$  and  $P6_3$ . For each space group, an automated search was performed for 5, 10, 20, 39 and 78 anomalous sites, which resulted in  $P6_1$  as the best space group among those tested up to this step and 19 anomalous sites. The following step was a series of ShelxE jobs where solvent content was varied between 0.25 and 0.75 in 0.05 increments, the best result was obtained with space group  $P6_5$  (the enantiomorph of  $P6_1$ ) for a solvent content of 0.45. The final solvent content calculated with MATTHEWS\_COEF was 76.2%. Datasets were collected on several additional crystals and data were analysed. One crystal was found to diffract to 2.15 Å when processed using the xia2/Dials pipeline<sup>27,28</sup>, followed by structure solution by molecular replacement with AMoRe (CCP4 package) using the model obtained from the first data set. A few residues belonging to turns or loops were built manually. The second data set was refined using Coot and RefMac, Phenix was used in the final steps and deposition. Final Ramachandran statistics showed 97.2% residues in favoured regions and 2.8% in allowed regions.

Chemical crosslinking and mass spectrometry. The samples subjected to chemical crosslinking were in 50 mM sodium phosphate buffer, pH 7.5, supplemented with 200 mM sodium chloride, 5% glycerol (vol/vol) and 10 mM  $\beta$ -mercaptoethanol. Samples were concentrated to 15 mg ml<sup>-1</sup>. The DSS crosslinker stock solution was 27.1 mM in DMF and the crosslinking reaction was carried out for two hours at room temperature, using 100x molar excess of DSS to protein sample. The samples were then reduced for 30 minutes at 60°C with 10 mM DTT in 50 mM ammonium bicarbonate buffer, alkylated for 30 minutes at room temperature using 50 mM iodoacetamide in ammonium bicarbonate buffer and digested with Sequencing Grade Modified Trypsin (Promega) for 17 hours at 37°C. The resulting peptide mixture was fractionated in five elution steps of 10, 15, 20, 30 and 80% acetonitrile (vol/vol) using Oasis HLB solid phase cartridges (Waters Corporation, USA) and loaded onto a nanoAcquity UPLC liquid chromatographer (Waters Corporation, USA) with an Acquity BEH C18 column (100  $\mu$ m x 100 mm, 1.7  $\mu$ m) coupled to a Waters Synapt HDMS mass spectrometer

(Waters Corporation, USA). 40 pmol of peptide mixture was analyzed in each run. The LC-MS/MS runs were processed on MASCOT Distiller v. 2.4 (Matrix Science Ltd) and analyzed by MASCOT v.2.4. (Matrix Science Ltd). These data obtained from MASCOT Distiller were then analyzed with Crux e SIM-XL to find the crosslinked peptides.

Hydrogen-deuterium exchange. Purified, full length TIPRL was incubated for 2 hours with the synthetic tetrapeptide DYFL or DpYFL (tyrosine phosphorylated) at a 1:10 protein:peptide molar ratio in 20 mM pH 7.5 Tris-HCl buffer, 100 mM NaCl, 10 mM  $\beta$ -mercaptoethanol. A control reaction was performed using dimethylformamide instead of the peptides. The hydrogen-deuterium exchange was started by diluting these reactions 1:15 in deuterated buffer at 25°C and stopped by adding equal volumes of quench buffer (800 mM Guanidine-HCl, 0.8% formic acid (vol/vol), 20 mM DTT 20, pH 2.5) at 4°C after 0s (control), 10s, 1 min, 10 min, 1h and 2h. Samples were then injected into a nano Acquity UPLC system with HDX technology coupled to Synapt G1 HDMS (Waters Corporation, USA). Online digestion was performed on an immobilized pepsin column (2x30 mm, Applied Biosystem, USA) for 3 minutes at 15°C with 50  $\mu$ L.min<sup>-1</sup> flow. The resulting peptides were desalted on an ACQUITY UPLC BEH C18 pre-column (1,7  $\mu$ m, VanGuard, Waters) at 0°C and separated on an analytical column (ACQUITY UPLC BEH C18 1,7  $\mu$ m, 1 mm x 100 mm, Waters) at 0°C with 50  $\mu$ L.min<sup>-1</sup> flow. The runs were processed by Protein Lynx Global Server v.2.4. (Waters Corporation, USA) and DynamX v.3.0 (Waters Corporation, USA)

GST pulldown assays. The expression plasmids for His-tagged, full length TIPRL and GST-tagged PP2Ac alpha, or the respective mutated constructs were cotransformed into *E. coli* BL21(DE3) and coexpression was induced by adding 0.5 mM IPTG and incubating for 5 hours at 25°C. The pellets were lysed using 20 mg ml<sup>-1</sup> lysozyme combined with sonication in pH 7.4 PBS buffer supplemented with 0.5% (v/v) IGEPAL CA-630, 1 mM DTT and 0.1mM PMSF. Glutathione sepharose beads (Glutathione-Sepharose 4B, GE Healthcare) were added to the cleared lysates, which were incubated for one hour at 4°C. The samples were washed three times with lysis

buffer and the beads were eluted with sample buffer and analyzed by coomassie stained SDS-PAGE.

### Acknowledgments

This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). V.S. was recipient of a FAPESP fellowship (2011/03054-9). R. A. was recipient of a CNPq productivity fellowship. We thank M. P. Dias, A. C. Alves and E. C. Teixeira for technical support and G. A. Bezerra for helpful insights. We thank LNBio/CNPEM for financial support and access to the facilities.

### AUTHOR CONTRIBUTIONS

J.H.C.S, V.S., R.A. and F.C.G. designed experiments and analyzed data. J.H.C.S. and N.I.T.Z designed the expression constructs and established the expression and purification protocols. V.S. purified and crystallized the protein and performed size exclusion and in vitro interaction assays. G.L.R. performed mutagenesis. V.S., R.A., J.B.N., J.S., and H.D.P. solved the crystal structure. T.L. and F.C.G. performed XL-MS and HDX experiments and analysis. J.H.C.S. and R.A. supervised the work and J.H.C.S. wrote the manuscript with input from other authors.

### References

1. Bryant, J. C., Westphal, R. S. & Wadzinski, B. E. Methylated C-terminal leucine residue of PP2A catalytic subunit is important for binding of regulatory Balpha subunit. *Biochemical Journal* **339**, 241 (1999).
2. Yu, X. X. *et al.* Methylation of the protein phosphatase 2A catalytic subunit is essential for association of Balpha regulatory subunit but not SG2NA, striatin, or polyomavirus middle tumor antigen. *Mol. Biol. Cell* **12**, 185–199 (2001).
3. Lee, J. & Lee, D.-H. Leucine methylation of protein phosphatase PP4C at C-terminal is critical for its cellular functions. *Biochem. Biophys. Res. Commun.* **452**, 42–47 (2014).
4. Shi, Y. Serine/threonine phosphatases: mechanism through structure. *Cell* **139**, 468–484 (2009).
5. Virshup, D. M. & Shenolikar, S. From promiscuity to precision: protein phosphatases get a makeover. *Mol. Cell* **33**, 537–545 (2009).
6. Jiang, L. *et al.* Structural basis of protein phosphatase 2A stable latency. *Nat Commun* **4**, 1699 (2013).



7. Guo, F. *et al.* Structural basis of PP2A activation by PTPA, an ATP-dependent activation chaperone. *Cell Res.* (2013). doi:10.1038/cr.2013.138
8. Stanevich, V. *et al.* The structural basis for tight control of PP2A methylation and function by LCMT-1. *Mol. Cell* **41**, 331–342 (2011).
9. Sents, W., Ivanova, E., Lambrecht, C., Haesen, D. & Janssens, V. The biogenesis of active protein phosphatase 2A holoenzymes: a tightly regulated process creating phosphatase specificity. *FEBS J.* **280**, 644–661 (2012).
10. Smetana, J. & Zanchin, N. I. T. Interaction analysis of the heterotrimer formed by the phosphatase 2A catalytic subunit, alpha4 and the mammalian ortholog of yeast Tip41 (TIPRL). *FEBS J.* **274**, 5891–5904 (2007).
11. Song, I. S. *et al.* Inhibition of MKK7-JNK by the TOR Signaling Pathway Regulator-Like Protein Contributes to Resistance of HCC Cells to TRAIL-Induced Apoptosis. *Gastroenterology* **143**, 1341–1351 (2012).
12. Jacinto, E., Guo, B., Arndt, K. T., Schmelzle, T. & Hall, M. N. TIP41 interacts with TAP42 and negatively regulates the TOR signaling pathway. *Mol. Cell* **8**, 1017–1026 (2001).
13. Nakashima, A. *et al.* A positive role of mammalian Tip41-like protein, TIPRL, in the amino-acid dependent mTORC1-signaling pathway through interaction with PP2A. *FEBS Lett.* (2013).
14. Leulliot, N. *et al.* Structure of protein phosphatase methyltransferase 1 (PPM1), a leucine carboxyl methyltransferase involved in the regulation of protein phosphatase 2A activity. *J. Biol. Chem.* **279**, 8351–8358 (2004).
15. Leulliot, N. *et al.* Crystal structure of the PP2A phosphatase activator: implications for its PP2A-specific PPLase activity. *Mol. Cell* **23**, 413–424 (2006).
16. Chao, Y. *et al.* Structure and mechanism of the phosphotyrosyl phosphatase activator. *Mol. Cell* **23**, 535–546 (2006).
17. Jordens, J. *et al.* The protein phosphatase 2A phosphatase activator is a novel peptidyl-prolyl cis/trans-isomerase. *J. Biol. Chem.* **281**, 6349–6357 (2006).
18. LeNoue-Newton, M. *et al.* The E3 ubiquitin ligase- and protein phosphatase 2A (PP2A)-binding domains of the Alpha4 protein are both required for Alpha4 to inhibit PP2A degradation. *J. Biol. Chem.* **286**, 17665–17671 (2011).
19. Xing, Y. *et al.* Structural mechanism of demethylation and inactivation of protein phosphatase 2A. *Cell* **133**, 154–163 (2008).
20. Xu, Y. *et al.* Structure of the protein phosphatase 2A holoenzyme. *Cell* **127**, 1239–1251 (2006).
21. Zhang, Y. I-TASSER: fully automated protein structure prediction in CASP8. *Proteins* **77 Suppl 9**, 100–113 (2009).
22. Matsuoka, S. *et al.* ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160–1166 (2007).
23. McConnell, J. L., Gomez, R. J., McCorvey, L. R. A., Law, B. K. & Wadzinski, B. E. Identification of a PP2A-interacting protein that functions as a negative regulator of phosphatase activity in the ATM/ATR signaling pathway. *Oncogene* **26**, 6021–6030 (2007).

24. Magnusdottir, A. *et al.* The crystal structure of a human PP2A phosphatase activator reveals a novel fold and highly conserved cleft implicated in protein-protein interactions. *J. Biol. Chem.* **281**, 22434–22438 (2006).
25. Löw, C. *et al.* Structural basis for PTPA interaction with the invariant C-terminal tail of PP2A. *Biol. Chem.* **395**, 881–889 (2014).
26. Yang, J., Roe, S. M., Prickett, T. D., Brautigan, D. L. & Barford, D. The structure of Tap42/alpha4 reveals a tetratricopeptide repeat-like fold and provides insights into PP2A regulation. *Biochemistry* **46**, 8807–8815 (2007).
27. Winter, G., Lobley, C. M. C. & Prince, S. M. Decision making in xia2. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1260–1273 (2013).

## FIGURE LEGENDS

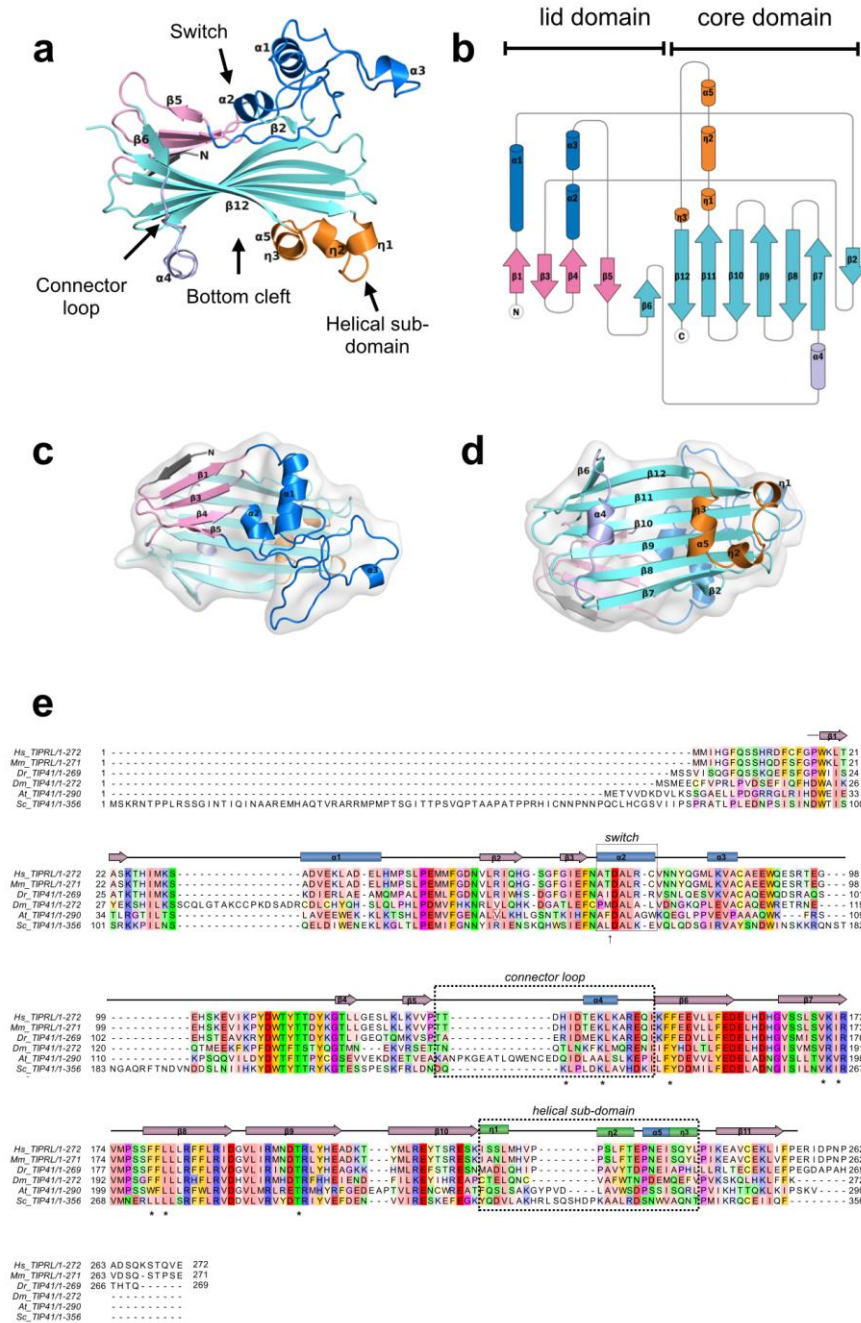


Figure 1. Overall structure of TIPRL. (a) Cartoon representation of the TIPRL fold highlighting its structural elements. (b) Topology diagram of the TIPRL fold generated with TopDraw. (c) Top and (d) bottom views of the cartoon representation combined with transparency representation of the surface where the secondary structure elements are labelled from N- to C-terminus. (e) The sequence conservation pattern of TIPRL/TIP41. The amino acid sequences of TIPRL from *Homo sapiens* (Hs) and *Mus musculus* (Mm) and TIP41 from *Danio rerio* (Dr), *Drosophila melanogaster* (Dm), *Arabidopsis thaliana* (At) and *Saccharomyces cerevisiae* (Sc) were aligned using ClustalW and color coded according to their degree of evolutionary conservation and physicochemical properties using the color scheme Zappo 21 from Jalview (aliphatic/hydrophobic: salmon, aromatic: orange, positive: blue, negative: red, hydrophilic: green, conformationally special: pink, cysteine: yellow). The secondary structure elements from the human TIPRLAN structure are represented on top of the alignment. Pink arrows represent beta strands, blue rectangles represent  $\alpha$ -helices and green rectangles represent  $3_{10}$  helices.

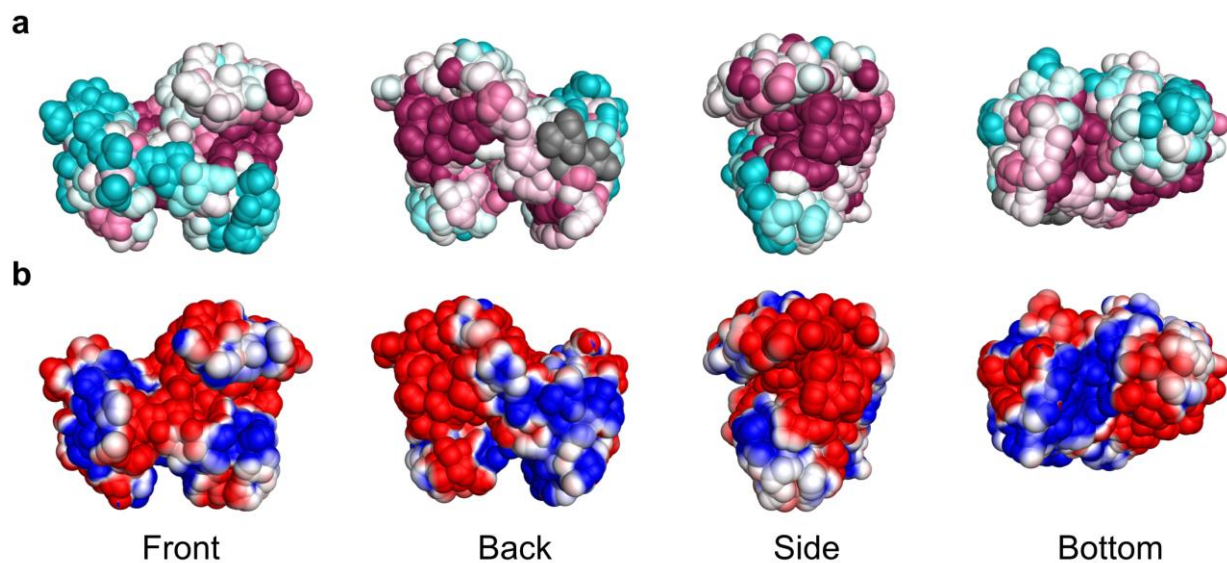


Figure 2. Surface properties of TIPRL. (a) The sequence conservation pattern of TIPRL was obtained using the Consurf server. The surface representation is colored from the least conserved (cyan) to the most conserved residues (magenta). (b) Surface representation of the macromolecular electrostatics calculation by the PyMol APBS (Adaptive Poisson-Boltzmann Solver).

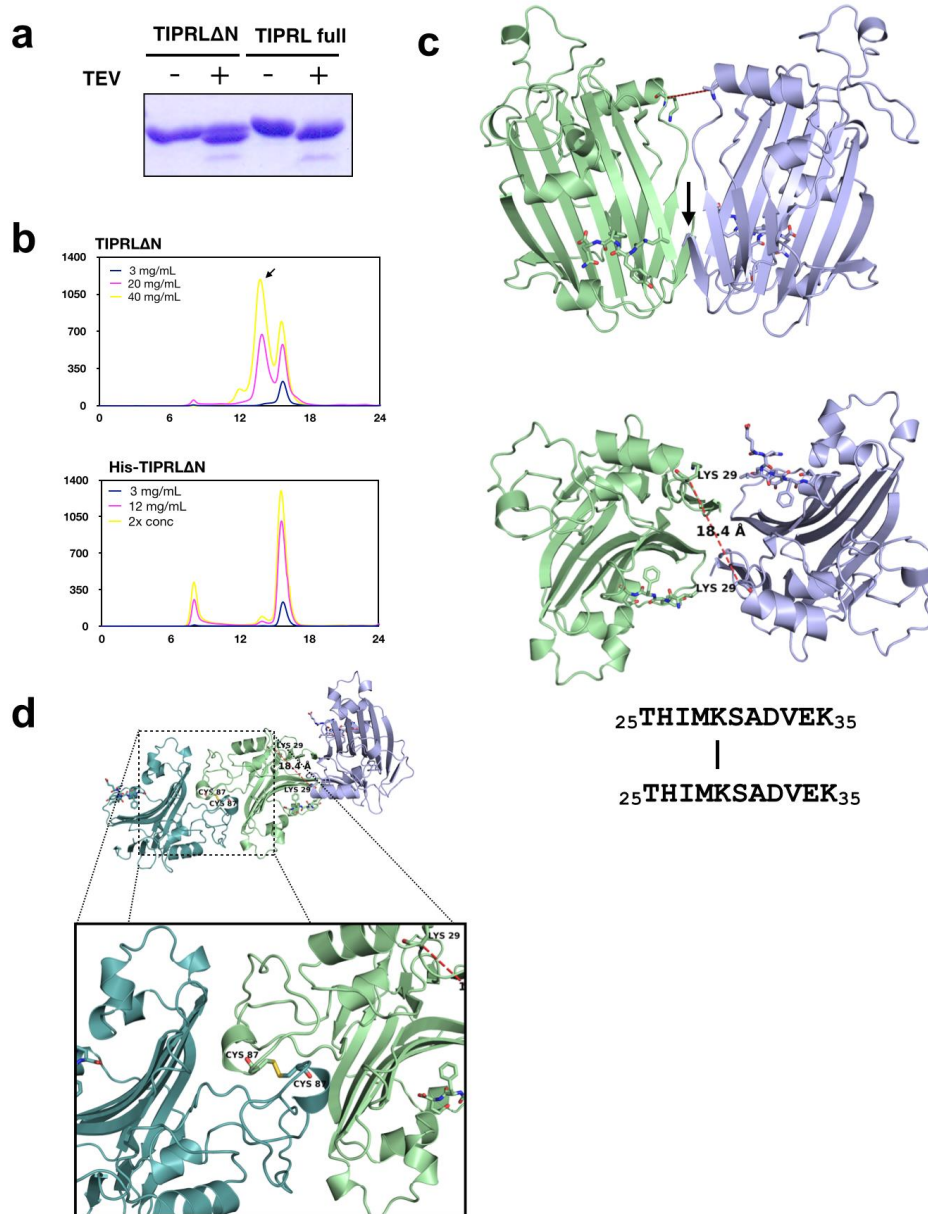


Figure 3. Dimerization of TIPRLΔN. (a) The coomassie stained SDS-PAGE analysis of purified TIPRLΔN and full length TIPRL, before and after removal of the histidine tag with TEV protease, shows a different cleavage pattern between these two constructs. TIPRLΔN undergoes partial cleavage, while full length TIPRL is completely cleaved under the same conditions. (b) Size exclusion profiles of TIPRLΔN on a Superdex 200 10/30 column, before or after cleavage with the TEV protease. TIPRLΔN (top panel) was analyzed at three different concentrations - 3, 20 and 40 mg ml<sup>-1</sup>, showing a concentration-dependent dimerization which was observed as a second peak eluting around 14 ml, indicated with a black arrow. Uncleaved TIPRLΔN (bottom panel), as well as full length TIPRL both cleaved and uncleaved (not shown) were not able to form this second peak under similar conditions. (c) The crystal structure of TIPRLΔN contains two molecules in the asymmetric unit, depicted as green and blue cartoon representations. The RMSD between these two molecules is 0.3. The dimer interface involves the N-terminus, more specifically the β1 strand. Crosslinking-MS analysis detected an interpeptide crosslink between Lys29 in both molecules, suggesting that this N-

terminal dimer is also observed in solution. Lys29 is shown in red and the 19.1 Å distance to its counterpart in the other molecule is indicated with a yellow trace. The black arrow indicates the position of the N-terminus. (d) A disulfide bond between Cys87 and its symmetrical counterpart stabilizes an interface between MolA (pale green) and MolB from the adjacent unit cell (teal).

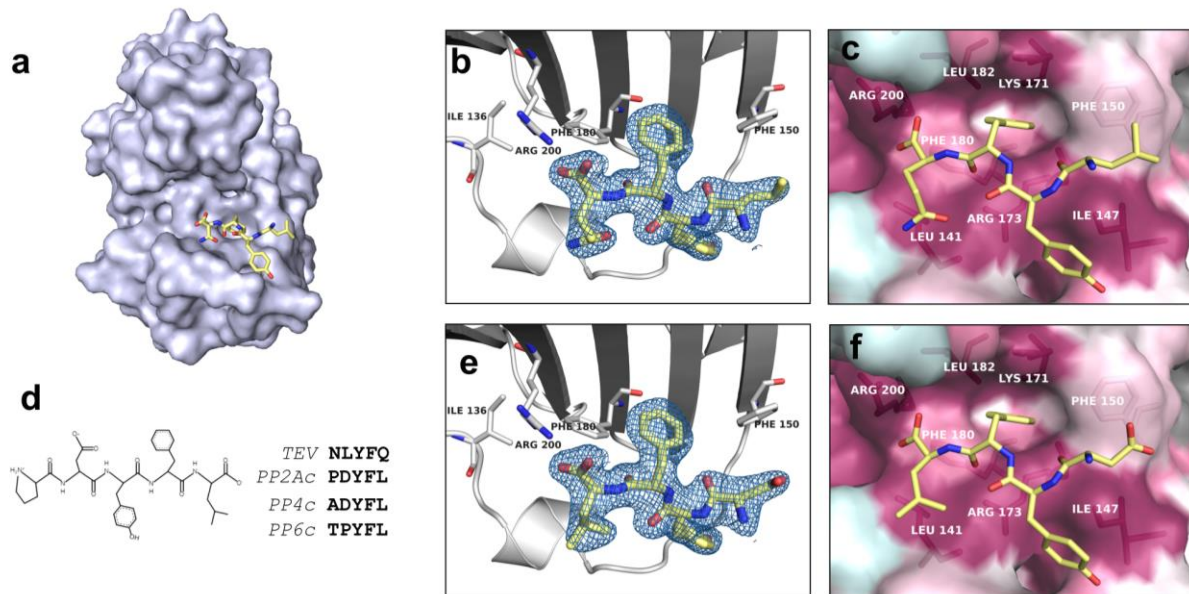


Figure 4. The bottom cleft of TIPRL harbours a conserved peptide binding site (a) Overview of the last four residues of the peptide generated after cleavage with the TEV protease (MGHHHHHHENLYFQ), LYFQ, bound to TIPRL (MolA). The peptide fragment is depicted as sticks over a surface representation of TIPRL, viewed from the bottom. (b) Close up view of the TEV site peptide LYFQ fitted into the electron density map depicted at  $3\sigma$ . (c) Conservation pattern of the peptide binding site showing the LYFQ peptide in sticks representation. Conservation was obtained from the Consurf server as in Figure 2a and the relevant amino acid side chains are labelled underneath the surface representation. (d) Sequence similarity between the TEV site peptide and the conserved C-terminus of type 2A phosphatases (PP2Ac, PP4c and PP6c). (e) The C-terminus of PP2Ac, DYFL, was fitted into the electron density map of the LYFQ peptide depicted at  $3\sigma$ . (f) Same representation as figure 4c, showing the PP2Ac-derived peptide DYFL instead of TEV peptide.

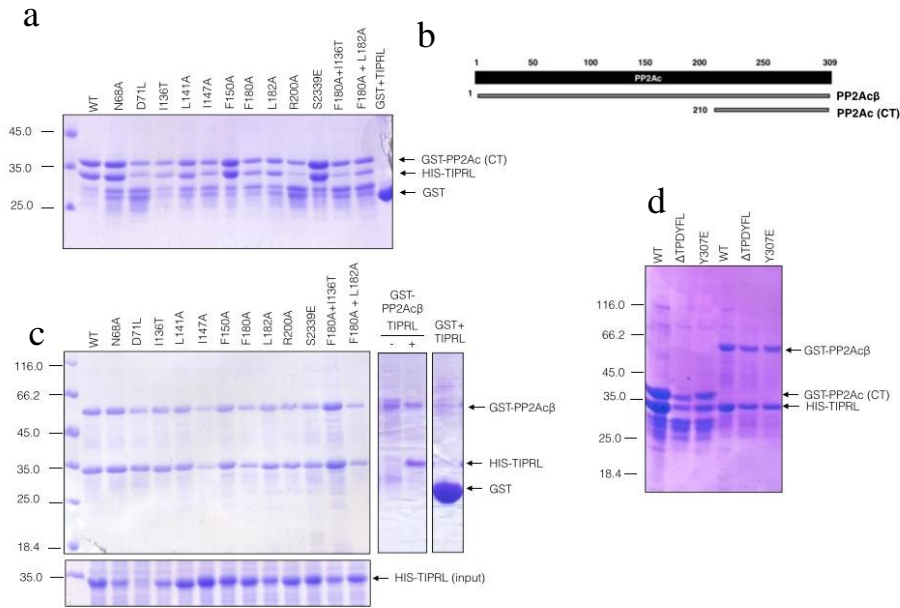


Figure 5. Validation of the PP2A C-terminus binding site on TIPRL by mutagenesis and pulldown. (a) GST-pulldown assays of mutated TIPRL. GST-tagged PP2Ac (residues 210 - 309) and His-tagged, full length TIPRL or its mutated versions were co-expressed in *E. coli* and the lysates were incubated with glutathione-sepharose beads. The mutants R200A and F180A + I136T show substantial loss of binding to PP2Ac $\alpha$ . (b) Schematic representation of the PP2A constructs used for GST pulldown. (c) The TIPRL mutants in (a) were assayed for their interaction with full length PP2Ac $\beta$ , but no substantial effects were observed for any of the mutations. The lower panel depicts the soluble lysates, showing reduced expression levels for D71L. (c) GST-pulldown assays of mutated PP2Ac $\alpha$  were performed as in (a), but in this case wild type TIPRL was combined with mutated versions of GST-tagged PP2Ac $\alpha$  (deletion of the four C-terminal residues of PP2A or mutation of Tyr307 to glutamate). Again, no appreciable reduction was observed.



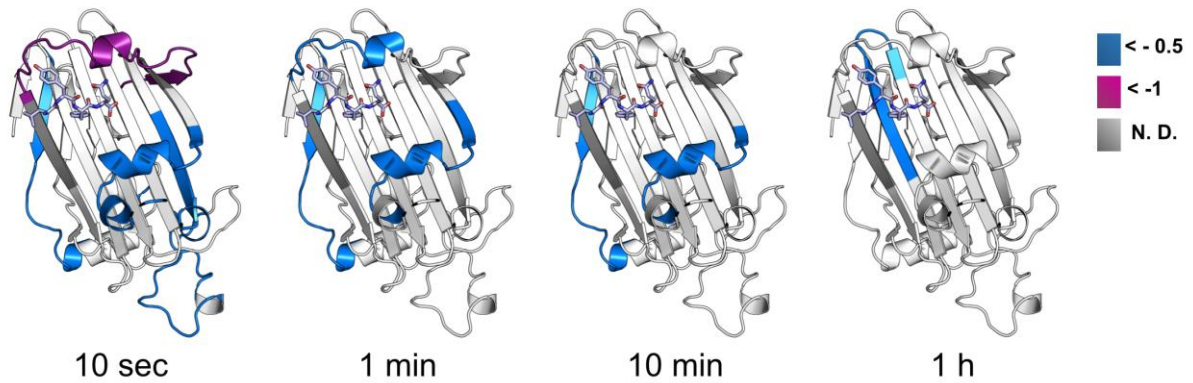
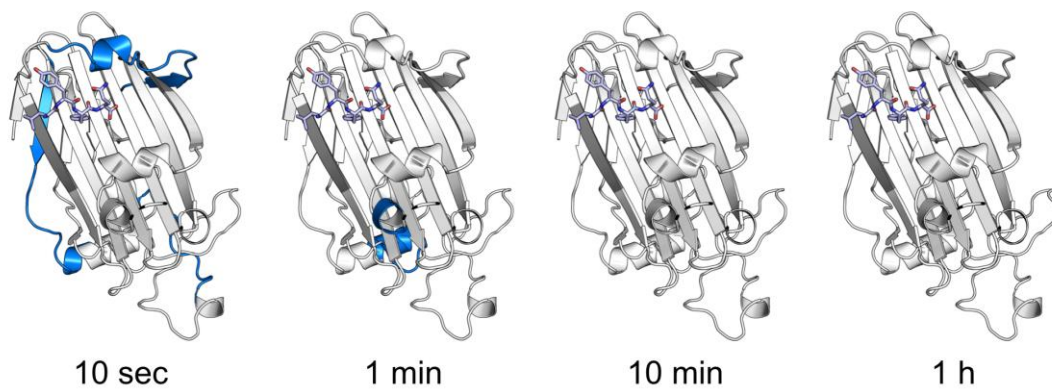
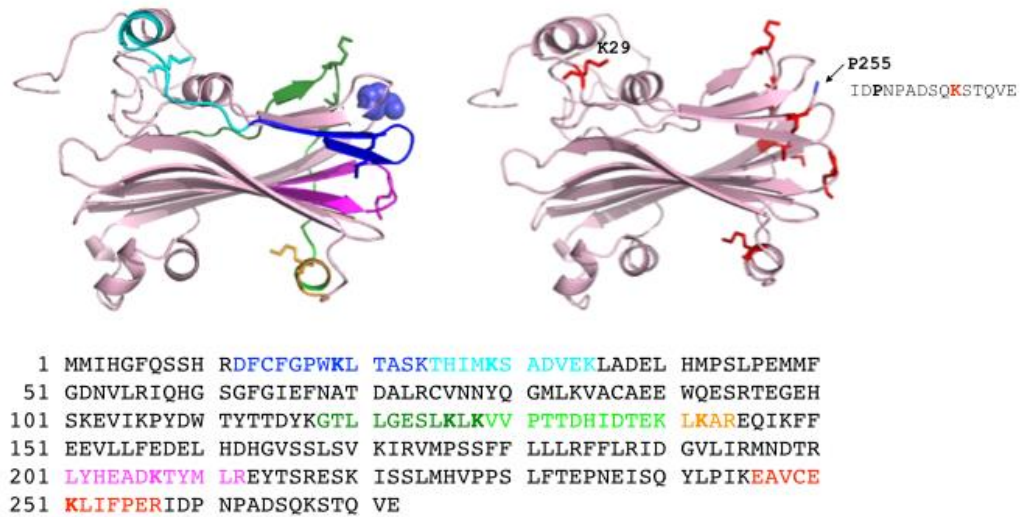
**a TIPRL+ DYFL****b TIPRL + DpYFL**

Figure 6. Structural representation of the HDX-MS analysis of full length, his-tagged TIPRL, in the presence of a synthetic tetrapeptide which mimics the PP2A C-terminus, unmodified (a) or phosphorylated (b). The bottom side of TIPRL is shown in cartoon representation and the DYFL peptide is represented as blue sticks. The relative deuterium exchange is represented as follows: Blue: 0 to -0.5 Da, Purple: 0.5 to -1.0 Da, Dark grey: not determined. In both cases, only negative values were observed, indicating either a gain of structure or reduced solvent exposure compared to the control sample with no bound peptide. The 2 hour time points are identical to the respective 1 hour timepoint and therefore are not depicted in the figure.



Supplementary Figure 1. Crosslinking/MS analysis of the C-terminus of TIPRL. The last residue visible in the electron density map was P255 (black arrow) , suggesting that residues 256 - 272 might be flexible. The c-terminal lysine 267 was crosslinked to each one of the lysines depicted as red sticks on the right side panel. Their tryptic peptides detected by MS are highlighted in different colors on the left side panel and the TIPRL sequence (bottom).

### Supplementary table 1. Sequences of mutagenic primers

Mutants	Sequence (5' - 3')
<b>TIPRL</b>	
N68A-F	catcttaacgcatctgtagcagcgaactcaattccaaagccaga
N68A-R	tctggctttggaattgagttcgcgtgctacagatgCGTTAAGATG
D71L-F	gttgtttacacatcttaacgctaagttagcattgaaactcaattccaaagccaga
D71L-R	tctggctttggaattgagttcaatgctacattagcgttaagatgtgtaacaac
I136T-F	ctctggctttcaatttttctgtatccgatatgatctggttaggtacaacctt
I136T-R	aaggttgtacctacaacagatcatacggatcacagaaaaattgaaagccagag
L141A-F	cttaatctggtctctggctttcgcgttttctgtatctatatgatctgtttag
L141A-R	ctacaacagatcatatagatacagaaaaagcgaagccagagaaacagattaag
I147A-F	aaaggagaacttcttcaaaaaacttagcctgttctctggctttcaatttttctg
I147A-R	cagaaaaattgaaagccagagaaacaggctaagtttttgaagaagttctccttt
F150A-F	aaaaggagaacttctcagcaaaacttaactgttctctggctttcaatttttctg
F150A-R	cagaaaaattgaaagccagagaaacagattaagtttgctgaagaagttctcctttt
F180A-F	ccgcaacagcagggcaagctagaaggcattactctaatac
F180A-R	gattagagtaatgccttctagctttgcctgtgctgttgcgg
L182A-F	ctcaagaaaaaccgcaacgcccaggaaaaagctagaaggca
L182A-R	tgcttctagcttttctggcgttgcggttttctttag
R184A-F	ccatcaattctcaagaaaaacgccaacagcagaaaaagctaga
R184A-R	tctagcttttctgctgttggcgttttcttgagaattgatgg
R200A-F	ttgtcagcctcatggtaaagtgcctgtgcattcattctgataag
R200A-R	cttatcagaatgaatgacacggcactttaccatgaggctgacaa
S239E-F	cttcctttattggtaaatactgctctatttcataggttccgtgaagaggggaag
S239E-R	cttccctttcacggaacctaatgaaatagagcagatattaccaataaagggaag
F180L182AA-F	tctcaagaaaaaccgcaacgcccaggcaaaagctagaaggcattactctaattctcaca
F180L182AA-R	tgtgaagattagagtaatgccttctagctttgcctggcgttgcggttttcttgaga
<b>PP2Ac</b>	
C-Terminal _Y307E_F	aggagaaatztataggaactcgtctggggtgcgcccgtg
C-Terminal _Y307E_R	cacggcgcaaccagacgagttcctataaatttctcct
C-Terminal _T304Stop	_Faatttataggaagtagtctggctagcgcctgtgtaacatgaggctc
C-Terminal _T304Stop	_Rgagcctcatgttacacggcgtagccagactacttccataaaatt
PP2Acβ _Y307E_F	attggatccttacaggaactcgtctggggtacgacgag
PP2Acβ _Y307E_R	ctcgtcgtacccagacgagttcctgtaaggatccaat
PP2Acβ _T304Stop_F	cttacaggaagtagtctggctaaccgacgagtaacatgtggctc
PP2Acβ _T304Stop_R	gagccacatgttactcgtcgttagccagactacttccctgtaag

# **CAPÍTULO 5**

## **Discussão**

## 5. Discussão

### 5.1 Discussão dos resultados dos estudos com PTP1B

A proteína tirosina fosfatase PTP1B é um importante alvo para o desenvolvimento de novos fármacos, já que sua superexpressão está associada a algumas doenças como o diabetes tipo 2 e o câncer. Devido a isso, houve um foco da indústria farmacêutica no desenvolvimento de moléculas inibidoras da tirosina fosfatase PTP1B, entretanto, sem sucesso. Os fatores limitantes para desenvolver bons inibidores da proteína tirosina fosfatase são a conservação evolutiva do sítio catalítico que reduz a especificidade dos compostos, já que as fosfatases da família proteína tirosina fosfatase têm o mesmo motivo catalítico, e a natureza química do sítio ativo altamente carregado com carga positiva. Sendo assim, as moléculas inibidoras também precisam ter essa característica, porém carregadas negativamente (Krishnan et al. 2014). Embora seja possível sintetizar inibidores carregados potentes, essa característica limita o desenvolvimento de drogas devido à dificuldade de atravessar a membrana plasmática das células, que é hidrofóbica.

Outra questão na dificuldade de desenvolver ligantes específicos para PTP1B é a promiscuidade de inibidores, como foi observado neste trabalho. A série de inibidores baseada no produto natural RK-682 foi triada para a proteína PTP1B e, em paralelo, para as outras proteínas tirosina fosfatases produzidas no grupo de pesquisa, as fosfatases de baixo peso LMW-PTP e a de especificidade dual CDC25B. Essa molécula inibiu de forma promíscua todas as proteínas tirosina fosfatases testadas. Esses estudos renderam a publicação Carneiro, Trivella *et. al.* 2015 intitulada: “*Is RK-682 a promiscuous enzyme inhibitor? Synthesis and in vitro evaluation of protein tyrosine phosphatase inhibition of racemic RK-682 and analogues*”. A autora desta tese também é uma das autoras deste artigo sobre a inibição das fosfatases por RK-682 (Carneiro *et. al.*, 2015).

Dentre as fosfatases, a proteína PTP1B é a mais suscetível à oxidação. Ela tem um estágio de formação do intermediário sulfenil-amida, o qual ocorre lentamente fazendo com que o *loop* catalítico de desloque para proteger a cisteína catalítica. O enxofre da cadeia lateral do resíduo de aminoácido C215 liga-se à

amina da cadeia principal da S216 num processo é reversível. Há outros três estágios de oxidação determinados pelo número de átomos de oxigênio ligados ao enxofre da cadeia lateral da C215, forma irreversível. Observou-se esse problema com recorrência. Muitas moléculas com ação inibidora estavam agindo por essa via e, assim, gerando falsos positivos na triagem de compostos. Esses compostos não podem ser usados no desenvolvimento de fármacos, pois liberam oxigênio no meio reacional, oxidando a cisteína catalítica. Existe uma frente de autores que apoiam o desenvolvimento de fármacos para o câncer baseado na liberação de oxigênio no meio (Kamb 2007). No entanto, há necessidade de mais pesquisas na área, pois o excesso de espécies oxidativas de oxigênio pode causar danos à célula.

Uma alternativa para ter melhores resultados na triagem de compostos e na cristalização da proteína com o ligante é o uso de ferramentas computacionais para um *screen* virtual e, somente depois das moléculas triadas virtualmente, realizar ensaios experimentais. Recentemente, a autora desta tese colaborou para uma publicação sobre racionalização de inibidores para a proteína fosfatase de baixo peso LMW-PTP, intitulada: “*Crystal structures of the apo form and a complex of human LMW-PTP with a phosphonic acid provide new evidence of a secondary site potentially related to the anchorage of natural substrates*”, na qual foram desenhados inibidores de LMW-PTP a partir da descoberta de uma molécula de PMSF (fenil-metil-sulfonil fluoreto) no sítio catalítico da enzima por cristalografia de proteínas (Fonseca *et. al.*, 2015).

Outra alternativa na busca de inibidores da PTP1B é o estudo da região de resíduos de aminoácidos 299-311, que é considerada um sítio alostérico. Essa região não é carregada e pouco conservada entre as fosfatases, e já tem sido explorada com bons resultados de inibição para o desenvolvimento de fármacos (Krishnan *et. al.*, 2014). Aqui, essa região não foi estudada, pois trabalhou-se com a forma truncada da proteína PTP1B no resíduo de aminoácido 298.

## 5.2 Discussão dos resultados dos estudos com TIPRL

### 5.2.1 Dimerização

A proteína TIPRL foi expressa e purificada na forma solúvel e em grande quantidade. A forma truncada, TIPRL $\Delta$ N, foi mais solúvel em relação à TIPRL *Full* e, em solução, forma um homodímero em condições muito específicas, tais como, proteólise da cauda de histidinas, concentrações acima de 10 mg.mL<sup>-1</sup> da amostra e deleção de 12 resíduos de aminoácidos no N-terminal. Provavelmente, na célula, a TIPRL assumia, unicamente, a forma de monômero, uma vez que TIPRL *Full* não forma dímeros e para a forma truncada dimerizar necessita daquelas condições. Sendo assim, pode-se considerar a dimerização como um artefato, o qual foi essencial para a formação dos cristais de proteína para os estudos cristalográficos, já que somente o dímero forma cristais e está presente como sendo o conteúdo da unidade assimétrica. A interface de dimerização envolve a região do N-terminal, como observado através do *cross-linking* MS, no qual foi identificada uma ligação cruzada entre o mesmo peptídeo (<sup>25</sup>THIMKSADVEK<sup>35</sup> - <sup>25</sup>THIMKSADVEK<sup>35</sup>). Na estrutura tridimensional, a presença da cauda de histidina e dos 12 resíduos bloquearia essa interface, ou seja, impediria a dimerização.

A região de interface do dímero é muito conservada e está estruturalmente próxima do sítio de ligação do C-terminal PP2Ac, ou seja, existe uma superfície contínua e conservada desde a interface de dimerização até a região que liga o C-terminal da PP2Ac. Isso sugere que a interface de dimerização no cristal seja a mesma que se liga ao domínio catalítico da PP2Ac e outras fosfatases 2A na célula.

O fato da TIPRL *Full* não dimerizar na célula reforça a ideia de que a PP2Ac ligue-se a essa porção. Curiosamente, a estrutura da Tap42, ortóloga de levedura da  $\alpha$ 4, foi relatada como um dímero em 2007 (Yang *et. al.*, 2007) e em estudos posteriores, observou-se que a interface de homodimerização no cristal de Tap42 é a mesma região de interação com a PP2Ac no cocrystal de  $\alpha$ 4 com PP2Ac (Jiang *et. al.*, 2013).

### 5.2.2 Obtenção da Estrutura

A determinação da estrutura de macromoléculas biológicas por cristalografia de Raios X envolve uma série de passos, como a seleção da molécula alvo, clonagem, expressão, purificação e cristalização, coleta de dados de difração e determinação de posições atômicas (Chayen e Saridakis 2008). No entanto, mesmo quando a proteína está solúvel e pura, produzir cristais de alta qualidade continua sendo um dos os principais entraves para a determinação da estrutura. E quando se tem dados de boa qualidade, outro obstáculo é a obtenção das fases para a determinação da estrutura.

O experimento de difração de Raios X fornece apenas a intensidade da onda espalhada em cada direção, e não a fase, o que leva ao problema da fase, já que é essencial saber esses dois parâmetros para o cálculo do mapa de densidade eletrônica e, desse modo, construir o modelo tridimensional (Chayen e Saridakis 2008). Dentre os métodos para resolver o Problema das Fases em Cristalografia, de modo a obter-se mapas de densidade eletrônica interpretáveis e construir um modelo para a proteína-alvo, encontram-se Métodos Diretos, Substituição Molecular e métodos baseados em sinal anômalo e/ou isomorfo (Taylor 2003; Rossmann e Arnold 2001; Blundell e Johnson 1976; Carter Jr. e Sweet 1997a; Carter Jr. e Sweet 1997b).

Os métodos diretos aplicam ferramentas matemáticas para obter a estrutura tridimensional e são empregados principalmente no caso de pequenas moléculas, onde a relação entre dados disponíveis e parâmetros é muito favorável.

A Substituição Molecular é uma técnica na qual se emprega um modelo de coordenadas atômicas obtido a partir de alguma proteína homóloga disponível no PDB (Protein Data Bank) ou modelos gerados computacionalmente. Esta técnica tem como principais limitações a identidade sequencial entre a proteína de interesse e a homóloga (quanto maior a identidade, maior as chances de sucesso) e a proporção em relação ao total de aminoácidos esperados (quanto maior a porção alinhada, melhor), além do número de monômeros contidos na unidade assimétrica (ASU) do cristal (quanto mais monômeros a serem encontrados, menores as chances de sucesso). Como a TIPRLAN não possui nenhuma homóloga cuja



estrutura esteja disponível no PDB ou outros bancos de dados e mesmo não havendo um modelo computacional ideal, a Substituição Molecular foi testada sem sucesso com modelos gerados pelos seguintes programas:

- I-TASSER (Zhang 2008; Roy *et. al.*, 2010; Roy *et. al.*, 2012) utiliza o alinhamento (*threading*) da sequência da proteína de interesse contra a biblioteca de estruturas do PDB e a partir disso gera cinco modelos com baixo *score*. Atualmente, é considerado o melhor programa de modelagem estrutural de proteínas.
- SAXSTER (dos Reis *et. al.*, 2011) similar ao anterior, porém, utilizando vínculos de SAXS durante o processo de *threading*.
- Rosetta (Rohl *et. al.*, 2004; Rohl 2005) monta pequenos fragmentos de proteínas conhecidas por uma estratégia de *Monte Carlo* para produzir as conformações da proteína de interesse. São gerados cinco modelos com mínimo de energia livre.
- QUARK (Xu e Zhang 2012) também usa apenas a sequência de aminoácido da proteína de interesse, porém só modela proteínas com até 200 resíduos. Este programa usa o algoritmo *ab initio* e simulações *Monte Carlo*. O programa é adequado para proteínas sem modelos homólogos como a TIPRLΔN, entretanto, quando usamos esse método precisamos excluir 72 resíduos de aminoácidos da sequência peptídica.
- BALBES (Long *et. al.*, 2008) diferente dos programas citados acima, ele utiliza o conjunto de dados de difração cristalográfica (arquivo *.mtz*), a sequência peptídica e a biblioteca do PDB para construir os modelos.

Os dois primeiros métodos de resolver a estrutura cristalográfica não foram viáveis para resolver a estrutura da TIPRLΔN, então, partiu-se para o último método, baseado em sinal anômalo e/ou isomorfo. Esses métodos consistem na inserção de átomos pesados dentro do cristal (Rossman e Arnold 2001; Blundell e Johnson 1976; Carter Jr. e Sweet 1997a; Carter Jr. e Sweet 1997b). Esses átomos causam desvios na Lei de Friedel,  $I(h,k,l) = I(-h,-k,-l)$ , de modo que a esperada igualdade não mais ocorre. Na presença de átomos pesados, as diferenças entre as

intensidades das reflexões  $I(h,k,l)$  e  $I(-h,-k,-l)$  são usadas como fonte de informação para a obtenção de fases iniciais. Nestes métodos, após melhoramento da densidade eletrônica, constrói-se um modelo atômico a partir dos mapas iniciais obtidos, seguindo-se, então, o refinamento estrutural.

A incorporação dos átomos de selênio foi a melhor opção para resolver o problema das fases dos cristais da proteína TIPRLAN, já que a incorporação de átomos pesados por cocristalização e *soaking* danificava os cristais e não era uma garantia da presença do átomo pesado. A proteína de interesse tem 10 metioninas em sua sequência peptídica de 272 resíduos de aminoácidos, resultando em uma razão de 27,2 resíduos de aminoácidos para cada selenometionina. A inserção da selenometionina garantiu a presença do espalhador anômalo. Então, em busca da perturbação nos pares de Friedel causada pelo espalhador, cristais da proteína TIPRLAN com a marcação de selenometionina foram coletados na linha de luz I02 do Diamond Light Source/Inglaterra. Com o uso de ferramentas computacionais, foi possível identificar a localização dos átomos de selênio e obter fases suficientemente boas para construir um modelo inicial da estrutura tridimensional. Este método de obter a estrutura não é muito usual no país. Os depósitos no banco de estrutura de macromoléculas (PDB) cresce continuamente, o que torna a substituição molecular a primeira opção.

### 5.2.3 Estrutura

A estrutura da TIPRLAN pertence a um *fold* novo, desvendou-se um novo tipo de enovelamento, fato extremamente raro atualmente, o último caso foi reportado em 2008. A estrutura tridimensional da TIPRL humana é primeira da família TIP41. Sendo assim, pode auxiliar a desvendar as estruturas das proteínas ortólogas que pertencem a essa família. A estrutura da TIP41, ortóloga de levedura, talvez, seja um pouco distinta da TIPRL humana; pois a de fungo pode formar um heterodímero com a Tap42, ortóloga da  $\alpha 4$ , diferente da humana que não forma esse complexo (Smetana e Zanchin 2007; McConnell *et. al.*, 2007; Jacinto *et. al.*, 2001).

Entretanto, a descoberta mais relevante da estrutura TIPRL $\Delta$ N foi mapear o sítio de interação com o C-terminal da proteína fosfatase PP2Ac. A região de interação é, de modo geral, hidrofóbica e, também, carregada positivamente, o que sugere ser uma região sem atividade de catálise enzimática. Os aminoácidos que compõem o sítio de ligação com a PP2Ac são bastante conservados nas outras espécies de eucariotos, como em leveduras, mostrando que a interação entre essas duas proteínas é muito antiga na linha evolutiva.

As abordagens anteriores para mapear o sítio de interação entre a TIPRL e a PP2Ac eram baseadas em mutagênese sítio-dirigidas em resíduos de aminoácidos aleatórios ou em experimentos baseados na isoforma menor da TIPRL, a TIP<sub>i</sub>2. A TIPRL de 20 kDa variante de *splicing*, tem 178 resíduos de aminoácidos enquanto a TIPRL tem 272 resíduos e não possui a mesma porção C-terminal da TIPRL. McConnell e colaboradores (McConnell *et. al.*, 2007) relataram que esta isoforma não interage com PP2Ac e, portanto, concluir que o local de ligação está localizado na região C-terminal de TIPRL (resíduos 173-272), o que corrobora os resultados obtidos ao longo desta tese. A região de interação entre a TIPRL e PP2Ac é pelo C-terminal de ambas. Os resíduos de aminoácidos I136, F180 e R200 da TIPRL são fundamentais para a interação com o C-terminal da PP2Ac, como apresentado nos resultados de *pull-down*, e a TIP<sub>i</sub>2 não possui esses resíduos. Além disso, a isoforma de 20 kDa não é enovelada corretamente, uma vez que só teria a primeira fita  $\beta$  que compõem a longa folha- $\beta$  que forma o C-terminal da TIPRL.

O mapeamento da região de interação com a PP2Ac se deu pela presença do peptídeo gerado a partir da clivagem com TEV que estava ligado à TIPRL, o qual pode ser observado no mapa densidade eletrônica da estrutura da TIPRL $\Delta$ N. A proteína PTPA, a qual adiciona o cátion bivalente à PP2Ac, foi cristalizada e em sua estrutura, também, apareceu peptídeo gerado da clivagem pela protease TEV (Magnusdottir *et. al.*, 2006). Os sete resíduos de aminoácidos foram encontrados em um bolsão hidrofóbico e foi interpretado como uma evidência para alguma função de ligação entre a PTPA e outra proteína, entretanto os autores não mencionam a semelhança com a PP2Ac C-terminal. Anos mais tarde, em 2014, observou por cristalografia que, na mesma região da PTPA, onde o peptídeo do sítio

de clivagem da TEV havia sido ligado na estrutura anterior, o C-terminal da PP2Ac interage no complexo PTPA-PP2Ac-ATPγS (Guo *et. al.*, 2014). Por conseguinte, é razoável argumentar que o peptídeo gerado a partir da clivagem com TEV reproduza a especificidade de ligação do C-terminal da PP2Ac. No caso específico aqui relatado, o local de ligação na TIPRL foi ainda validado com mutagênese sítio-dirigida, *pull-down* e experimento de troca de hidrogênio por deutério (HDX). Essa região de interação da TIPRL com a PP2Ac, também, é a mesma região de interação da TIPRL com as fosfatases PP4 e PP6, uma vez que a sequência DYFL do C-terminal é idêntica ao da PP2Ac.

#### 5.2.4 Provável função biológica da TIPRL

A biogênese da PP2A é um processo complexo e pouco compreendido. O que se propõe na literatura atualmente é que PP2Ac forma um complexo com  $\alpha 4$ , no qual a fosfatase está protegida da degradação proteolítica (Kong *et. al.*, 2009; Jiang *et. al.*, 2013). Em seguida ocorre a incorporação do cátion bivalente pela PTPA (Cayla *et. al.*, 1994) seguido da carboximetilação no resíduo de aminoácido L309 do extremo C-terminal pela LCMT1 (De Baere *et. al.*, 1999) e, em algum momento, a fosforilação do resíduo Y307 por uma quinase (Chen *et. al.*, 1992). Em uma recente revisão, Sents e colaboradores (Sents *et. al.*, 2012) propõem que cinco proteínas são responsáveis pela biogênese de PP2Ac:  $\alpha 4$ , PTPA, LCMT1 e PME-1 e, potencialmente, TIPRL. Durante a última década, vários estudos elucidaram os mecanismos de regulação PP2Ac por estas proteínas através de suas estruturas isoladas ou em complexo com PP2Ac. No entanto, os estudos do papel biológico da TIPRL permaneceram desconhecidos, salvo por alguns estudos que indicam que a interação entre ela e a PP2Ac causava a inibição da atividade catalítica *in vitro* da enzima (Smetana e Zanchin 2007), provavelmente a inibição seja através do bloqueio do sítio ativo da PP2Ac causado pela TIPRL, assim como ocorre com as outras 4 proteínas que participam da biogênese.

Em 2013, (Nakashima *et. al.*, 2013) mostrou-se que a TIPRL liga-se, preferencialmente, à PP2Ac metilada e não quando fosforilada em Y307. Então, aqui

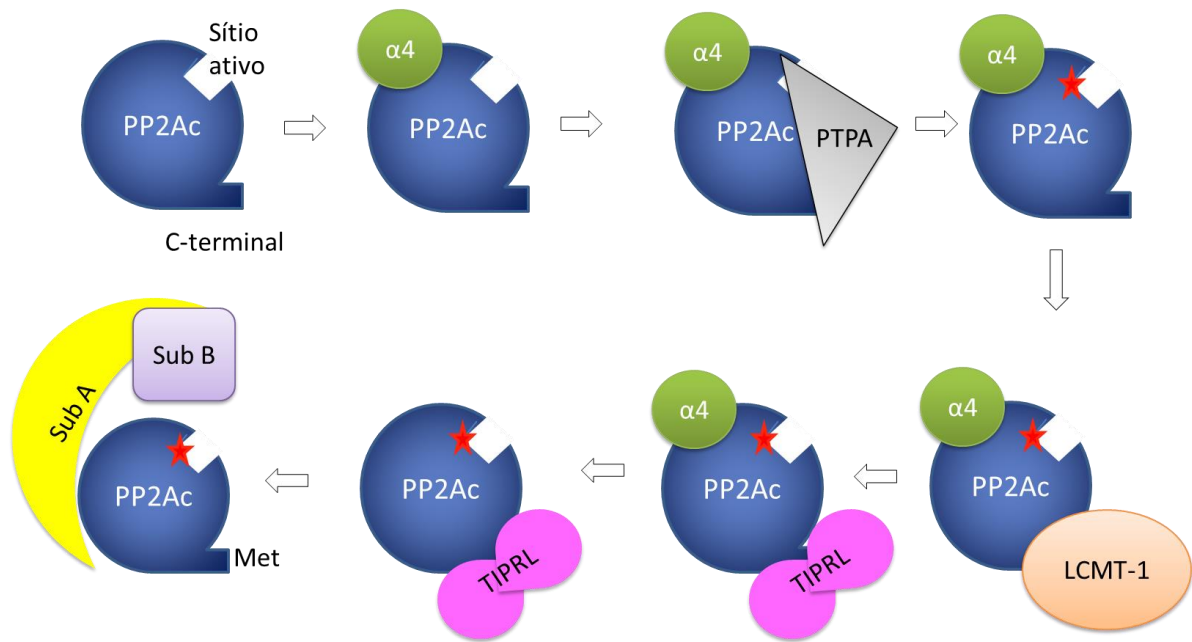
é proposto que a TIPRL liga-se à PP2Ac depois da LCMT1 e antes da quinase adicionar o fosfato no resíduo de tirosina. Na estrutura cristalográfica da TIPRL foi possível observar que no bolsão de interação com o C-terminal da PP2Ac há espaço para o encaixe de um grupo metil no extremo C-terminal, o qual pode interagir com os resíduos hidrofóbicos do fundo do sítio. Já para o fosfato, não há espaço disponível e este é carregado negativamente o que dificulta a interação com as porções hidrofóbicas do sítio. Além disso, o fosfato é mais volumoso que o grupo metil, o que atrapalha o encaixe dessa modificação pós-traducional no bolsão da TIPRL. Os dados de HDX desta tese indicam que a TIPRL é mais estruturada quando ligada ao peptídeo sintético do C-terminal da PP2Ac.

Aqui é proposto que a PP2Ac liga-se à TIPRL depois da LMCT1, de modo que a formação desse complexo estabiliza ambas as proteínas. A TIPRL se molda para encaixar e estabilizar a PP2Ac, já que no experimento de HDX com os peptídeos sintéticos, o peptídeo não fosforilado causou mudanças de estabilidade em toda a proteína, não apenas no bolsão hidrofóbico de ligação sendo que a porção N-terminal da TIPRL passou a ser mais estável, provavelmente. Há contatos extensos entre TIPRL com a PP2Ac, pois no experimento de *pull-down* com a PP2Ac truncada no resíduo T304, sem os últimos 5 resíduos que interagem com o bolsão da TIPRL, a interação entre elas foi mantida e, também, a mutação no resíduo D71 da TIPRL *in vivo* faz com que ocorra a perda de interação entre elas (Smetana e Zanchin 2007; Nakashima *et. al.*, 2013). No entanto, *in vitro*, não foi observado perda da interação entre a PP2Ac e TIPRL com mutante D71, diferente do que relatado em células humanas.

Os resíduos de aminoácidos mutados na fenda onde o C-terminal da PP2Ac se encaixa na TIPRL não alteram a interação entre a TIPRL e a PP2Ac *Full*. De modo que há diferença na interação entre a TIPRL mutada e os cem resíduos de aminoácidos que compõem a menor região de interação com a PP2Ac, o C-terminal (resíduos 210-309). Os mutantes R200A e I136A/F180A, que compõem o sítio, reduzem a interação entre a TIPRL e o C-terminal da PP2Ac truncada.

O complexo TIPRL:PP2Ac se mantém até a ligação da PP2Ac com as suas subunidades de ancoragem A e regulatória B para a formação do heterotrímero funcional. Quando as subunidades estão ligadas a PP2Ac, a TIPRL não interage

com a catalítica. A TIPRL pode atuar como o freio na biogênese da PP2A, segurando e mantendo estável a PP2Ac junto com a  $\alpha 4$ . Talvez, o complexo PP2Ac:TIPRL: $\alpha 4$  receba alguma sinalização química, a qual faz com que a  $\alpha 4$  seja liberada e a TIPRL encaminhe a PP2Ac para a montagem do heterotrímero, porém não se sabe em qual momento isso pode ocorrer. (Figura 5.1).



**Figura 5.1:** Provável função da TIPRL na biogênese da PP2A. A TIPRL atuaria com um freio na montagem do heterotrímero funcional, depois da LMCT1 e seria o último estágio para a forma ativa da fosfatase. A estrela em vermelho representa o cátion bivalente.

A função de freio da TIPRL na biogênese da PP2A pode ser explorada como um possível alvo para o desenvolvimento de drogas anticâncer, uma vez que a TIPRL inibe a PP2A, que é um supressor tumoral, e quando a primeira estiver com o sítio de interação ocupado por pequena molécula, a segunda pode ser ativada. Estudos recentes mostram que ao inibir a TIPRL por siRNA em hepatocarcinoma humano, o tumor regride (Nakashima *et. al.*, 2013). Além disso, o gene *knockout* da TIPRL em embriões de *zebrafish* promoveu apoptoses (Song *et. al.*, 2012). Yoon e colaboradores, 2015, (Yoon *et. al.*, 2015) observaram que o complexo entre TIPRL e MKK7 é quebrado na presença de frações que continham ácido chicórico, o qual é, estruturalmente, semelhante à porção C-terminal da PP2Ac que interage com o

bolsão hidrofóbico da TIPRL . No entanto, as amostras do complexo TIPRL:PP2Ac tratados com esse ácido se mantiveram estáveis. Provavelmente, a TIPRL tenha uma maior afinidade a PP2Ac em relação a MKK7. Todavia, esse dado não invalida a possibilidade da TIPRL ser um alvo para o desenvolvimento de fármacos.

# **CAPÍTULO 6**

## **Conclusões e Perspectivas**



## 6. Conclusões Perspectivas

Os estudos de inibição com a proteína PTP1B identificaram uma boa quantidade de compostos com ação inibitória sobre a enzima, porém a maioria deles liberaram espécies reativas de oxigênio no meio reacional, o que causava a inibição da enzima por oxidação do resíduo catalítico ou não foram observados na densidade eletrônica. Apesar de não ter encontrado ligantes nas estruturas da PTP1B resolvidas ao longo da pesquisa, os trabalhos de triagem em colaborações com outros membros do grupo de pesquisa, renderam duas publicações, a primeira sobre a promiscuidade do composto RK-682 e a segunda reportando novas estruturas de LMW-PTP apo e com ligantes.

O comportamento distinto entre a TIPRL *Full* e TIPRL $\Delta$ N em relação à dimerização, sugere que, na célula, a TIPRL esteja na forma monomérica, o que é uma novidade, pois essa informação não existia antes deste trabalho. Os resíduos de aminoácidos que participam da interface de dimerização são conservados evolutivamente entre as proteínas TIPRL de outros organismos e é a mesma região onde, possivelmente, a PP2Ac seja ancorada *in vivo*, informação que pode contribuir para o estudo do complexo PP2Ac:TIPRL, futuramente.

Neste trabalho, para obter a estrutura tridimensional inédita da proteína TIPRL $\Delta$ N foram utilizadas técnicas, tais como a Substituição Molecular e o *cross-linking* acoplado a MS, sem sucesso. Esse problema só foi solucionado empregando a técnica de SAD (dispersão anômala a um único comprimento de onda), através da incorporação da selenometionina na cadeia peptídica da proteína (técnica pouco usual no país), e o envio dos cristais para a coleta na linha de luz do Diamond Light Source/Inglaterra. Com a estrutura tridimensional resolvida, constatou-se que se tratava de um novo enovelamento, o que é raro atualmente. Essa descoberta poderá contribuir na resolução da estrutura das outras proteínas que compõem a família Tip41.

A possível função biológica da TIPRL só pôde ser proposta através da estrutura tridimensional com o ligante, peptídeo gerado a partir da clivagem com a protease TEV, e estudos de *pull-down* associados a mutagêneses. Antes deste trabalho, pouquíssimo era conhecido sobre essa proteína. As descobertas desta

tese abrem caminho tanto para um entendimento do mecanismo molecular de regulação das fosfatases 2A por TIPRL, quanto para abordagens racionais para explorar a interface de interação.

As interfaces de interação proteína-proteína são de grande interesse para a indústria farmacêutica, porém, estas ainda são pouco exploradas devido à dificuldade de obter pequenas moléculas para atingir na maioria dessas interfaces, que em geral são planares. Apenas 12 moduladores de interface proteína-proteína para o tratamento do câncer estão em testes clínicos ou em uso atualmente. Alguns casos especiais de interfaces se prestam mais facilmente como alvo de drogas - por exemplo, uma  $\alpha$ -hélice dentro de uma fenda (“*helix-in-groove*”). Estas são chamadas de “*druggable interfaces*” e se assemelham estruturalmente aos clássicos “*druggable pockets*” em sítios ativos de enzimas (Nero *et. al.*, 2014). Como perspectivas, a autora desta tese buscará futuramente obter pequenas moléculas que interagem na região de interação entre a TIPRL e as proteínas fosfatases 2A e diminuam ou desfaçam a interação.

## 7. Referências

- Barford, D., A. J. Flint, and N. K. Tonks. 1994. Crystal structure of human protein tyrosine phosphatase 1B. *Science* 263 (5152):1397-1404.
- Bentires-Alj, M., and B. G. Neel. 2007. Protein-tyrosine phosphatase 1B is required for HER2/Neu-induced breast cancer. *Cancer Res* 67 (6):2420-2424.
- Bleasdale, J. E., D. Ogg, B. J. Palazuk, C. S. Jacob, M. L. Swanson, X. Y. Wang, D. P. Thompson, R. A. Conradi, W. R. Mathews, A. L. Laborde, C. W. Stuchly, A. Heijbel, K. Bergdahl, C. A. Bannow, C. W. Smith, C. Svensson, C. Liljebris, H. J. Schostarez, P. D. May, F. C. Stevens, and S. D. Larsen. 2001. Small molecule peptidomimetics containing a novel phosphotyrosine bioisostere inhibit protein tyrosine phosphatase 1B and augment insulin action. *Biochemistry* 40 (19):5642-5654.
- Blundell, T. L., and L. N. Johnson. 1976. *Protein Crystallography*. New York, USA.: Academic Press.
- Buhrman, G., B. Parker, J. Sohn, J. Rudolph, and C. Mattos. 2005. Structural mechanism of oxidative regulation of the phosphatase Cdc25B via an intramolecular disulfide bond. *Biochemistry* 44 (14):5307-5316.
- Carneiro, V. M., D. B. Trivella, V. Scorsato, V. L. Beraldo, M. P. Dias, T. J. Sobreira, R. Aparicio, and R. A. Pilli. 2015. Is RK-682 a promiscuous enzyme inhibitor? Synthesis and in vitro evaluation of protein tyrosine phosphatase inhibition of racemic RK-682 and analogues. *Eur J Med Chem* 97:42-54.
- Carter Jr., C. W., and R. M. Sweet. 1997a. *Macromolecular Crystallography, Part A, volume 276 of Methods in Enzymology*. San Diego.: Academic Press.
- Carter Jr., C. W., and R. M. Sweet. 1997b. *Macromolecular Crystallography, Part B, volume 277 of Methods in Enzymology*. San Diego: Academic Press.
- Cayla, X., C. Van Hoof, M. Bosch, E. Waelkens, J. Vandekerckhove, B. Peeters, W. Merlevede, and J. Goris. 1994. Molecular cloning, expression, and characterization of PTPA, a protein that activates the tyrosyl phosphatase activity of protein phosphatase 2A. *J Biol Chem* 269 (22):15668-15675.
- Chayen, N. E., and E. Saridakis. 2008. Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods* 5 (2):147-153.

- Chen, J., B. L. Martin, and D. L. Brautigan. 1992. Regulation of protein serine-threonine phosphatase type-2A by tyrosine phosphorylation. *Science* 257 (5074):1261-1264.
- Cho, U. S., and W. Xu. 2007. Crystal structure of a protein phosphatase 2A heterotrimeric holoenzyme. *Nature* 445 (7123):53-57.
- Cudney, R., S. Patel, K. Weisgraber, Y. Newhouse, and A. McPherson. 1994. Screening and optimization strategies for macromolecular crystal growth. *Acta Crystallogr D Biol Crystallogr* 50 (Pt 4):414-423.
- De Baere, I., R. Derua, V. Janssens, C. Van Hoof, E. Waelkens, W. Merlevede, and J. Goris. 1999. Purification of porcine brain protein phosphatase 2A leucine carboxyl methyltransferase and cloning of the human homologue. *Biochemistry* 38 (50):16539-16547.
- dos Reis, M. A., R. Aparicio, and Y. Zhang. 2011. Improving Protein Template Recognition by Using Small-Angle X-Ray Scattering Profiles. *Biophysical Journal* 101 (11):2770-2781.
- Elchebly, M., P. Payette, E. Michaliszyn, W. Cromlish, S. Collins, A. L. Loy, D. Normandin, A. Cheng, J. Himms-Hagen, C. C. Chan, C. Ramachandran, M. J. Gresser, M. L. Tremblay, and B. P. Kennedy. 1999. Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science* 283 (5407):1544-1548.
- Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60 (Pt 12 Pt 1):2126-2132.
- Ewald, P. P., ed. 1962. *Fifty years of X-Ray diffraction*: IUCr.
- Fauman, E. B., J. P. Cogswell, B. Lovejoy, W. J. Rocque, W. Holmes, V. G. Montana, H. Piwnica-Worms, M. J. Rink, and M. A. Saper. 1998. Crystal structure of the catalytic domain of the human cell cycle control phosphatase, Cdc25A. *Cell* 93 (4):617-625.
- Fioramonte, M., A. M. dos Santos, S. McIlwain, W. S. Noble, K. G. Franchini, and F. C. Gozzo. 2012. Analysis of secondary structure in proteins by chemical cross-linking coupled to MS. *Proteomics* 12 (17):2746-2752.
- Fonseca, E. M. 2015. LMW-PTP, CDC25B, Inibidores de fosfatases implicadas em câncer , Química farmacêutica , Planejamento racional de fármacos ,

- Cristalografia de proteínas. 16/01/2015. 160 folhas. Tese de doutorado. Universidade Estadual de Campinas/ Campinas SP 16/01/2015.
- Fonseca, E. M., D. B. Trivella, V. Scorsato, M. P. Dias, N. L. Bazzo, K. R. Mandapati, F. L. de Oliveira, C. V. Ferreira-Halder, R. A. Pilli, P. C. Miranda, and R. Aparicio. 2015. Crystal structures of the apo form and a complex of human LMW-PTP with a phosphonic acid provide new evidence of a secondary site potentially related to the anchorage of natural substrates. *Bioorg Med Chem*.
- Franklin, R. E., and R. G. Gosling. 1953. Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature* 172 (4369):156-157.
- Gildea, R. J., D. G. Waterman, J. M. Parkhurst, D. Axford, G. Sutton, D. I. Stuart, N. K. Sauter, G. Evans, and G. Winter. 2014. New methods for indexing multi-lattice diffraction data. *Acta Crystallographica Section D-Biological Crystallography* 70:2652-2666.
- Glusker, J. P. 1994. Dorothy Crowfoot Hodgkin (1910-1994). *Protein Sci* 3 (12):2465-2469.
- Goldberg, J., H. B. Huang, Y. G. Kwon, P. Greengard, A. C. Nairn, and J. Kuriyan. 1995. Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature* 376 (6543):745-753.
- Guimaraes, B. G., L. Sanfelici, R. T. Neuenschwander, F. Rodrigues, W. C. Grizolli, M. A. Raulik, J. R. Piton, B. C. Meyer, A. S. Nascimento, and I. Polikarpov. 2009. The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS. *J Synchrotron Radiat* 16 (Pt 1):69-75.
- Guo, F., V. Stanevich, N. Wlodarchak, R. Sengupta, L. Jiang, K. A. Satyshur, and Y. Xing. 2014. Structural basis of PP2A activation by PTPA, an ATP-dependent activation chaperone. *Cell Res* 24 (2):190-203.
- Haesen, D., W. Sents, K. Lemaire, Y. Hoorne, and V. Janssens. 2014. The Basic Biology of PP2A in Hematologic Cells and Malignancies. *Front Oncol* 4:347.
- Han, Y. M., H. Oh, M. Na, B. S. Kim, W. K. Oh, B. Y. Kim, D. G. Jeong, S. E. Ryu, D. E. Sok, and J. S. Ahn. 2005. PTP1B inhibitory effect of abietane diterpenes isolated from *Salvia miltiorrhiza*. *Biol Pharm Bull* 28 (9):1795-1797.
- Hanahan, D., and R. A. Weinberg. 2000. The hallmarks of cancer. *Cell* 100 (1):57-70.
- Hanahan, D., and R. A. Weinberg. 2011. Hallmarks of cancer: the next generation. *Cell* 144 (5):646-674.

- Heneberg, P. 2009. Use of protein tyrosine phosphatase inhibitors as promising targeted therapeutic drugs. *Curr Med Chem* 16 (6):706-733.
- Hoppe, E., P. F. Berne, D. Stock, J. S. Rasmussen, N. P. Moller, A. Ullrich, and R. Huber. 1994. Expression, purification and crystallization of human phosphotyrosine phosphatase 1B. *Eur J Biochem* 223 (3):1069-1077.
- Hornbeck, P. V., J. M. Kornhauser, S. Tkachev, B. Zhang, E. Skrzypek, B. Murray, V. Latham, and M. Sullivan. 2012. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res* 40 (Database issue):D261-270.
- Hunter, T. 1995. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* 80 (2):225-236.
- INCA. 2015. *Instituto Nacional do Câncer - O que é o câncer?* . [http://www1.inca.gov.br/conteudo\\_view.asp?id=322](http://www1.inca.gov.br/conteudo_view.asp?id=322) 2015 [cited 17/07 2015].
- Incardona, M. F., G. P. Bourenkov, K. Levik, R. A. Pieritz, A. N. Popov, and O. Svensson. 2009. EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. *J Synchrotron Radiat* 16 (Pt 6):872-879.
- Jabbour, E., and H. Kantarjian. 2014. Chronic myeloid leukemia: 2014 update on diagnosis, monitoring, and management. *Am J Hematol* 89 (5):547-556.
- Jacinto, E., B. Guo, K. T. Arndt, T. Schmelzle, and M. N. Hall. 2001. TIP41 interacts with TAP42 and negatively regulates the TOR signaling pathway. *Mol Cell* 8 (5):1017-1026.
- Janssens, V., and J. Goris. 2001. Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem J* 353 (Pt 3):417-439.
- Janssens, V., J. Goris, and C. Van Hoof. 2005. PP2A: the expected tumor suppressor. *Curr Opin Genet Dev* 15 (1):34-41.
- Janssens, V., and A. Rebollo. 2012. The role and therapeutic potential of Ser/Thr phosphatase PP2A in apoptotic signalling networks in human cancer cells. *Curr Mol Med* 12 (3):268-287.

- Jiang, L., V. Stanevich, K. A. Satyshur, M. Kong, G. R. Watkins, B. E. Wadzinski, R. Sengupta, and Y. Xing. 2013. Structural basis of protein phosphatase 2A stable latency. *Nat Commun* 4:1699.
- Julien, S. G., N. Dube, M. Read, J. Penney, M. Paquet, Y. Han, B. P. Kennedy, W. J. Muller, and M. L. Tremblay. 2007. Protein tyrosine phosphatase 1B deficiency or inhibition delays ErbB2-induced mammary tumorigenesis and protects from lung metastasis. *Nat Genet* 39 (3):338-346.
- Junqueira, L. C. J., and J. Carneiro. 2005. *Biologia Celular e Molecular Guanabara Koogan* 8°:350.
- Kamb, A. W., S. Lengauer, C. 2007. Why is cancer drug discovery so difficult? *Nature Reviews Drug Discovery* 6:115-120.
- Kantardjieff, K., Rupp, B. 2003. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.* 12 (9):1865-71.
- Keegan, R. M., and M. D. Winn. 2008. MrBUMP: an automated pipeline for molecular replacement. *Acta Crystallogr D Biol Crystallogr* 64 (Pt 1):119-124.
- Kendrew, J. C., G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips. 1958. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* 181 (4610):662-666.
- Kennelly, P. J. 2003. Archaeal protein kinases and protein phosphatases: insights from genomics and biochemistry. *Biochem J* 370 (Pt 2):373-389.
- Klaman, L. D., O. Boss, O. D. Peroni, J. K. Kim, J. L. Martino, J. M. Zabolotny, N. Moghal, M. Lubkin, Y. B. Kim, A. H. Sharpe, A. Stricker-Krongrad, G. I. Shulman, B. G. Neel, and B. B. Kahn. 2000. Increased energy expenditure, decreased adiposity, and tissue-specific insulin sensitivity in protein-tyrosine phosphatase 1B-deficient mice. *Mol Cell Biol* 20 (15):5479-5489.
- Kong, M., D. Ditsworth, T. Lindsten, and C. B. Thompson. 2009. Alpha4 is an essential regulator of PP2A phosphatase activity. *Mol Cell* 36 (1):51-60.
- Krishnan, N., D. Koveal, D. H. Miller, B. Xue, S. D. Akshinthala, J. Kragelj, M. R. Jensen, C. M. Gauss, R. Page, M. Blackledge, S. K. Muthuswamy, W. Peti, and N. K. Tonks. 2014. Targeting the disordered C terminus of PTP1B with an allosteric inhibitor. *Nat Chem Biol* 10 (7):558-566.

- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227 (5259):680-685.
- Lambrecht, C., D. Haesen, W. Sents, E. Ivanova, and V. Janssens. 2013. Structure, regulation, and pharmacological modulation of PP2A phosphatases. *Methods Mol Biol* 1053:283-305.
- Langer, G. G., S. Hazledine, T. Wiegels, C. Carolan, and V. S. Lamzin. 2013. Visual automated macromolecular model building. *Acta Crystallogr D Biol Crystallogr* 69 (Pt 4):635-641.
- Larsen, S. D., T. Barf, C. Liljebris, P. D. May, D. Ogg, T. J. O'Sullivan, B. J. Palazuk, H. J. Schostarez, F. C. Stevens, and J. E. Bleasdale. 2002. Synthesis and biological activity of a novel class of small molecular weight peptidomimetic competitive inhibitors of protein tyrosine phosphatase 1B. *J Med Chem* 45 (3):598-622.
- Larsen, S. D., F. C. Stevens, T. J. Lindberg, P. M. Bodnar, T. J. O'Sullivan, H. J. Schostarez, B. J. Palazuk, and J. E. Bleasdale. 2003. Modification of the N-terminus of peptidomimetic protein tyrosine phosphatase 1B (PTP1B) inhibitors: identification of analogues with cellular activity. *Bioorg Med Chem Lett* 13 (5):971-975.
- Lee, S., and Q. Wang. 2007. Recent development of small molecular specific inhibitor of protein tyrosine phosphatase 1B. *Med Res Rev* 27 (4):553-573.
- Leslie, A. G. 1999. Integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* 55 (Pt 10):1696-1702.
- Leslie, A. G. 2006. The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* 62 (Pt 1):48-57.
- Li, W., Y. Zhang, and J. Skolnick. 2004. Application of sparse NMR restraints to large-scale protein structure prediction. *Biophys J* 87 (2):1241-1248.
- Liljebris, C., S. D. Larsen, D. Ogg, B. J. Palazuk, and J. E. Bleasdale. 2002. Investigation of potential bioisosteric replacements for the carboxyl groups of peptidomimetic inhibitors of protein tyrosine phosphatase 1B: identification of a tetrazole-containing inhibitor with cellular activity. *J Med Chem* 45 (9):1785-1798.



- Long, F., A. A. Vagin, P. Young, and G. N. Murshudov. 2008. BALBES: a molecular-replacement pipeline. *Acta Crystallographica Section D-Biological Crystallography* 64:125-132.
- Magnusdottir, A., P. Stenmark, S. Flodin, T. Nyman, M. Hammarstrom, M. Ehn, H. M. Bakali, H. Berglund, and P. Nordlund. 2006. The crystal structure of a human PP2A phosphatase activator reveals a novel fold and highly conserved cleft implicated in protein-protein interactions. *J Biol Chem* 281 (32):22434-22438.
- Manning, G., D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam. 2002. The protein kinase complement of the human genome. *Science* 298 (5600):1912-1934.
- Matsuoka, S., B. A. Ballif, A. Smogorzewska, E. R. McDonald, K. E. Hurov, J. Luo, C. E. Bakalarski, Z. M. Zhao, N. Solimini, Y. Lerenthal, Y. Shiloh, S. P. Gygi, and S. J. Elledge. 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 316 (5828):1160-1166.
- Matthews, B. W.. 1968. Solvent content of protein crystals. *J Mol Biol.* 28;33 (2):491-7.
- McConnell, J. L., R. J. Gomez, L. R. McCorvey, B. K. Law, and B. E. Wadzinski. 2007. Identification of a PP2A-interacting protein that functions as a negative regulator of phosphatase activity in the ATM/ATR signaling pathway. *Oncogene* 26 (41):6021-6030.
- McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read. 2007. Phaser crystallographic software. *J Appl Crystallogr* 40 (Pt 4):658-674.
- Migueleti, D. L., J. H. Smetana, H. F. Nunes, J. Kobarg, and N. I. Zanchin. 2012. Identification and characterization of an alternatively spliced isoform of the human protein phosphatase 2Aalpha catalytic subunit. *J Biol Chem* 287 (7):4853-4862.
- Monteleone, M. C., A. E. Gonzalez Wusener, J. E. Burdisso, C. Conde, A. Caceres, and C. O. Arregui. 2012. ER-bound protein tyrosine phosphatase PTP1B interacts with Src at the plasma membrane/substrate interface. *PLoS One* 7 (6):e38948.
- Moy, J. P., A. P. Hammersley, S. O. Svensson, A. Thompson, K. Brown, L. Claustre, A. Gonzalez, and S. McSweeney. 1996. A novel technique for accurate

- intensity calibration of area X-ray detectors at almost arbitrary energy. *Journal of Synchrotron Radiation* 3:1-5.
- Murshudov, G. N., A. A. Vagin, and E. J. Dodson. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53 (Pt 3):240-255.
- Nakashima, A., K. Tanimura-Ito, N. Oshiro, S. Eguchi, T. Miyamoto, A. Momonami, S. Kamada, K. Yonezawa, and U. Kikkawa. 2013. A positive role of mammalian Tip41-like protein, TIPRL, in the amino-acid dependent mTORC1-signaling pathway through interaction with PP2A. *FEBS Lett* 587 (18):2924-2929.
- Nero, T. L., C. J. Morton, J. K. Holien, J. Wielens, and M. W. Parker. 2014. Oncogenic protein interfaces: small molecules, big challenges. *Nature Reviews Cancer* 14 (4):248-262.
- Nobelprize. 2015. *Structure and function of the ribosome* 2009 [cited 17/07 2015].
- Olsen, J. V., B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, and M. Mann. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127 (3):635-648.
- Ostman, A., C. Hellberg, and F. D. Bohmer. 2006. Protein-tyrosine phosphatases and cancer. *Nat Rev Cancer* 6 (4):307-320.
- Paramelle, D., G. Miralles, G. Subra, and J. Martinez. 2013. Chemical cross-linkers for protein structure studies by mass spectrometry. *Proteomics* 13 (3-4):438-456.
- Pereira, S. R., V. M. Vasconcelos, and A. Antunes. 2011. The phosphoprotein phosphatase family of Ser/Thr phosphatases as principal targets of naturally occurring toxins. *Crit Rev Toxicol* 41 (2):83-110.
- Putnam, C. D., M. Hammel, G. L. Hura, and J. A. Tainer. 2007. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics* 40 (3):191-285.
- Reynolds, R. A., A. W. Yem, C. L. Wolfe, M. R. Deibel, Jr., C. G. Chidester, and K. D. Watenpaugh. 1999. Crystal structure of the catalytic subunit of Cdc25B required for G2/M phase transition of the cell cycle. *J Mol Biol* 293 (3):559-568.

- Rohl, C. A. 2005. Protein structure estimation from minimal restraints using rosetta. *Nuclear Magnetic Resonance of Biological Macromolecules, Part C* 394:244-260.
- Rohl, C. A., C. E. M. Strauss, K. M. S. Misura, and D. Baker. 2004. Protein structure prediction using rosetta. *Numerical Computer Methods, Pt D* 383:66-+.
- Rossmann, M. G., and E. Arnold. 2001. *International Tables for Crystallography, Vol. F. Crystallography of Biological Macromolecules*: Dordrecht: Kluwer Academic Publishers.
- Roy, A., A. Kucukural, and Y. Zhang. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5 (4):725-738.
- Roy, A., J. Y. Yang, and Y. Zhang. 2012. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Research* 40 (W1):W471-W477.
- Sacco, F., L. Perfetto, L. Castagnoli, and G. Cesareni. 2012. The human phosphatase interactome: An intricate family portrait. *FEBS Lett* 586 (17):2732-2739.
- Salton, S. R. 2005. Teaching resources. Protein phosphatases. *Sci STKE* 2005 (273):tr8.
- Santos, A. M., D. Schechtman, A. C. Cardoso, C. F. Clemente, J. C. Silva, M. Fioramonte, M. B. Pereira, T. M. Marin, P. S. Oliveira, A. C. Figueira, S. H. Oliveira, I. L. Torriani, F. C. Gozzo, J. Xavier Neto, and K. G. Franchini. 2011. FERM domain interaction with myosin negatively regulates FAK in cardiomyocyte hypertrophy. *Nat Chem Biol* 8 (1):102-110.
- Sents, W., E. Ivanova, C. Lambrecht, D. Haesen, and V. Janssens. 2012a. The biogenesis of active protein phosphatase 2A holoenzymes: a tightly regulated process creating phosphatase specificity. *FEBS J* 280 (2):644-661
- Seshacharyulu, P., P. Pandey, K. Datta, and S. K. Batra. 2013. Phosphatase: PP2A structural importance, regulation and its aberrant expression in cancer. *Cancer Lett*.
- Shapiro, A. L., E. Vinuela, and J. V. Maizel, Jr. 1967. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochem Biophys Res Commun* 28 (5):815-820.
- CELL\_NOW. University of Göttingen.

- Sheldrick, G. M. 2010. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr* 66 (Pt 4):479-485.
- Shi, L., H. P. Yu, Y. Y. Zhou, J. Q. Du, Q. Shen, J. Y. Li, and J. Li. 2008. Discovery of a novel competitive inhibitor of PTP1B by high-throughput screening. *Acta Pharmacol Sin* 29 (2):278-284.
- Shi, Y. 2009. Serine/threonine phosphatases: mechanism through structure. *Cell* 139 (3):468-484.
- Sjolin, L. 1993. 3D-structural elucidation of biologically important macromolecules. *Drug Des Discov* 9 (3-4):261-276.
- Smetana, J. H., C. L. Oliveira, W. Jablonka, T. Aguiar Pertinhez, F. R. Carneiro, M. Montero-Lomeli, I. Torriani, and N. I. Zanchin. 2006. Low resolution structure of the human alpha4 protein (IgBP1) and studies on the stability of alpha4 and of its yeast ortholog Tap42. *Biochim Biophys Acta* 1764 (4):724-734.
- Smetana, J. H., and N. I. Zanchin. 2007. Interaction analysis of the heterotrimer formed by the phosphatase 2A catalytic subunit, alpha4 and the mammalian ortholog of yeast Tip41 (TIPRL). *FEBS J* 274 (22):5891-5904.
- Song, I. S., S. Y. Jun, H. J. Na, H. T. Kim, S. Y. Jung, G. H. Ha, Y. H. Park, L. Z. Long, D. Y. Yu, J. M. Kim, J. H. Kim, J. H. Ko, C. H. Kim, and N. S. Kim. 2012. Inhibition of MKK7-JNK by the TOR signaling pathway regulator-like protein contributes to resistance of HCC cells to TRAIL-induced apoptosis. *Gastroenterology* 143 (5):1341-1351.
- Stanevich, V., L. Jiang, K. A. Satyshur, Y. Li, P. D. Jeffrey, Z. Li, P. Menden, M. F. Semmelhack, and Y. Xing. 2011. The structural basis for tight control of PP2A methylation and function by LCMT-1. *Mol Cell* 41 (3):331-342.
- Studier, F. W. 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41 (1):207-234.
- Svergun, D. I. 1992. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *Journal of Applied Crystallography* 25:495-503.
- Taylor, G. 2003. The phase problem. *Acta Crystallographica Section D-Biological Crystallography* 59:1881-1890.

- Terentyev, D., A. E. Belevych, R. Terentyeva, M. M. Martin, G. E. Malana, D. E. Kuhn, M. Abdellatif, D. S. Feldman, T. S. Elton, and S. Gyorke. 2009. miR-1 overexpression enhances Ca(2+) release and promotes cardiac arrhythmogenesis by targeting PP2A regulatory subunit B56alpha and causing CaMKII-dependent hyperphosphorylation of RyR2. *Circ Res* 104 (4):514-521.
- The CCP4 suite: programs for protein crystallography. 1994. *Acta Crystallogr D Biol Crystallogr* 50 (Pt 5):760-763
- Tonks, N. K. 2013. Protein tyrosine phosphatases--from housekeeping enzymes to master regulators of signal transduction. *FEBS J* 280 (2):346-378.
- Torrent, L., and I. Ferrer. 2012. PP2A and Alzheimer disease. *Curr Alzheimer Res* 9 (2):248-256.
- Uchida, K., P. G. Leiman, F. Arisaka, and S. Kanamaru. 2013. Structure and properties of the C-terminal beta-helical domain of VgrG protein from *Escherichia coli* O157. *J Biochem*.
- van Montfort, R. L., M. Congreve, D. Tisi, R. Carr, and H. Jhoti. 2003. Oxidation state of the active-site cysteine in protein tyrosine phosphatase 1B. *Nature* 423 (6941):773-777.
- Virshup, D. M. 2000. Protein phosphatase 2A: a panoply of enzymes. *Curr Opin Cell Biol* 12 (2):180-185.
- Watson, J. D., and F. H. Crick. 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171 (4356):737-738.
- Weinberg, R. A. 2008. *A biologia do câncer*. Edited by Artmed. Porto Alegre.
- Winter, G., C. M. Lobley, and S. M. Prince. 2013. Decision making in xia2. *Acta Crystallogr D Biol Crystallogr* 69 (Pt 7):1260-1273.
- Winter, G., and K. E. McAuley. 2011. Automated data collection for macromolecular crystallography. *Methods* 55 (1):81-93.
- Wong, L. L., D. Zhang, C. F. Chang, and E. S. Koay. 2010. Silencing of the PP2A catalytic subunit causes HER-2/neu positive breast cancer cells to undergo apoptosis. *Exp Cell Res* 316 (20):3387-3396.
- Xing, Y., Z. Li, Y. Chen, J. B. Stock, P. D. Jeffrey, and Y. Shi. 2008. Structural mechanism of demethylation and inactivation of protein phosphatase 2A. *Cell* 133 (1):154-163.

- Xing, Y., Y. Xu, Y. Chen, P. D. Jeffrey, Y. Chao, Z. Lin, Z. Li, S. Strack, J. B. Stock, and Y. Shi. 2006. Structure of protein phosphatase 2A core enzyme bound to tumor-inducing toxins. *Cell* 127 (2):341-353.
- Xu, D., and Y. Zhang. 2012. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins-Structure Function and Bioinformatics* 80 (7):1715-1735.
- Xu, Y., Y. Chen, P. Zhang, P. D. Jeffrey, and Y. Shi. 2008. Structure of a protein phosphatase 2A holoenzyme: insights into B55-mediated Tau dephosphorylation. *Mol Cell* 31 (6):873-885.
- Xu, Y., Y. Xing, Y. Chen, Y. Chao, Z. Lin, E. Fan, J. W. Yu, S. Strack, P. D. Jeffrey, and Y. Shi. 2006. Structure of the protein phosphatase 2A holoenzyme. *Cell* 127 (6):1239-1251.
- Yang, J., S. M. Roe, T. D. Prickett, D. L. Brautigan, and D. Barford. 2007. The structure of Tap42/alpha4 reveals a tetratricopeptide repeat-like fold and provides insights into PP2A regulation. *Biochemistry* 46 (30):8807-8815.
- Yang, W., W. A. Hendrickson, R. J. Crouch, and Y. Satow. 1990a. Structure of Ribonuclease-H Phased at 2-Å Resolution by Mad Analysis of the Selenomethionyl Protein. *Science* 249 (4975):1398-1405.
- Yang, W., W. A. Hendrickson, E. T. Kalman, and R. J. Crouch. 1990b. Expression, Purification, and Crystallization of Natural and Selenomethionyl Recombinant Ribonuclease-H from Escherichia-Coli. *Journal of Biological Chemistry* 265 (23):13553-13559.
- Yoon, J. Y., H. S. Cho, J. J. Lee, H. J. Lee, S. Y. Jun, J. H. Lee, H. H. Song, S. Choi, V. Saloura, C. G. Park, C. H. Kim, and N. S. Kim. 2015. Novel TRAIL sensitizer Taraxacum officinale F.H. Wigg enhances TRAIL-induced apoptosis in Huh7 cells. *Mol Carcinog*.
- Z. Otwinowski, W. M. 1997. Processing of X-ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymology*, edited by R. M. S. W. Carter Jr. New York: Academic Press, 307-326.
- Zabell, A. P., S. Corden, P. Helquist, C. V. Stauffacher, and O. Wiest. 2004. Inhibition studies with rationally designed inhibitors of the human low molecular weight protein tyrosine phosphatase. *Bioorg Med Chem* 12 (8):1867-1880.

- Zabell, A. P., A. D. Schroff, Jr., B. E. Bain, R. L. Van Etten, O. Wiest, and C. V. Stauffacher. 2006. Crystal structure of the human B-form low molecular weight phosphotyrosyl phosphatase at 1.6-Å resolution. *J Biol Chem* 281 (10):6520-6527.
- Zhang, M., C. V. Stauffacher, D. Lin, and R. L. Van Etten. 1998. Crystal structure of a human low molecular weight phosphotyrosyl phosphatase. Implications for substrate specificity. *J Biol Chem* 273 (34):21714-21720.
- Zhang, Y. 2008. I-TASSER server for protein 3D structure prediction. *Bmc Bioinformatics* 9.
- Zhang, Z. Y. 2002. Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development. *Annu Rev Pharmacol Toxicol* 42:209-234.
- Zhou, H., H. Singh, Z. D. Parsons, S. M. Lewis, S. Bhattacharya, D. R. Seiner, J. N. LaButti, T. J. Reilly, J. J. Tanner, and K. S. Gates. 2011. The biological buffer bicarbonate/CO<sub>2</sub> potentiates H<sub>2</sub>O<sub>2</sub>-mediated inactivation of protein tyrosine phosphatases. *J Am Chem Soc* 133 (40):15803-15805.

**Profa. Dra. Rachel Meneguello**  
Presidente  
Comissão Central de Pós-Graduação  
**Declaração**

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **Regulação de proteínas fosfatases envolvidas em câncer: estudos estruturais da TIPRL e interação com a Proteína Fosfatase 2A**, não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 15 de julho de 2015

Assinatura : Valéria Scorsato  
Nome do(a) autor(a): **Valéria Scorsato**  
RG n.º 41.310.175-7

Assinatura : RA  
Nome do(a) orientador(a): **Ricardo Aparicio**  
RG n.º 20.269.992-4



## DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação de Mestrado/tese de Doutorado intitulada Regulação de proteínas fosfatases envolvidas em câncer: estudos estruturais da TIPRL e interação com a Proteína Fosfatase 2A:

não se enquadra no § 4º do Artigo 1º da Informação CCPG 002/13, referente a bioética e biossegurança.

Tem autorização da(s) seguinte(s) Comissão(ões):

CIBio – Comissão Interna de Biossegurança , projeto No. RIAp 2011-2, Instituição: Instituto de Química/ UNICAMP.

CEUA – Comissão de Ética no Uso de Animais , projeto No. \_\_\_\_\_ , Instituição: \_\_\_\_\_

CEP - Comissão de Ética em Pesquisa, protocolo No. \_\_\_\_\_ , Instituição: \_\_\_\_\_

*\* Caso a Comissão seja externa ao IB/UNICAMP, anexar o comprovante de autorização dada ao trabalho. Se a autorização não tiver sido dada diretamente ao trabalho de tese ou dissertação, deverá ser anexado também um comprovante do vínculo do trabalho do aluno com o que constar no documento de autorização apresentado.*

Assinatura: Valéria Scorsato  
Nome do(a) aluno(a): Valéria Scorsato

Assinatura: RA  
Nome do(a) orientador(a): Prof. Dr. Ricardo Aparicio

Para uso da Comissão ou Comitê pertinente:

( ) Deferido ( ) Indeferido

Carimbo e assinatura

Para uso da Comissão ou Comitê pertinente:

( ) Deferido ( ) Indeferido

Carimbo e assinatura

Campinas, 15 de julho de 2015.

## DECLARAÇÃO

Declaramos para os devidos fins que o projeto de tese de doutorado intitulado: "Caracterização Estrutural Das Fosfatases Humanas PTP-1B E PP2A, Envolvidas Em Câncer, E Suas Interações Com Inibidores"; teve seu título alterado para: "Regulação de proteínas fosfatases envolvidas em câncer: estudos estruturais da TIPRL e interação com a Proteína Fosfatase 2A"; sendo este o título da Tese de Doutorado da aluna Valéria Scorsato.

Valéria Scorsato

Valéria Scorsato

Aluna de doutorado do PPGGBM

RA

Prof. Dr. Ricardo Aparicio

Orientador



COMISSÃO INTERNA DE BIOSSEGURANÇA  
UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE QUÍMICA

Cx. Postal 6154 - Campinas - 13083-970 SP - BRASIL  
<http://www.iqm.unicamp.br>



Solicitação/Protocolo: RiAp 2011-2

Nome do aluno: Valéria Scorsato.

Nível: doutorado

Período: março/2011 a agosto/2015

Nome do orientador: Ricardo Aparício

Titulo: Caracterização Estrutural Das Fosfatases Humanas PTP-1B E PP2A, Envolvidas Em Câncer, E Suas Interações Com Inibidores

Resumo: O câncer é uma das principais causas de morte atualmente, sendo a segunda causa de morte não acidental no Brasil, onde os gastos governamentais com essa patologia ultrapassam a casa de 1 bilhão de reais anualmente. Os tratamentos e terapias utilizados são falhos, sendo de extrema importância o desenvolvimento de novas terapias ou melhoramento das terapias já existentes. O desenvolvimento de moduladores da atividade de proteínas superexpressas em tumores e/ou proteínas que regulam etapas chave nos processos de captação de nutrientes e produção de energia, sinalização, divisão ou apoptose, reparo do DNA é uma das principais áreas de pesquisa nesta direção. Este projeto visa caracterizar bioquímica e estruturalmente, proteínas fosfatases humanas e proteínas envolvidas em vias de biogênese dessas fosfatases em suas formas recombinantes expressas em *E. coli*. As proteínas estudadas neste projeto foram identificadas como importantes alvos para o estudo do câncer, ou por serem superexpressas em tumores, ou por regularem vias de sinalização envolvidas no reparo dos danos ao DNA. Os dados a serem obtidos contribuirão para o entendimento das bases moleculares da interação dessas enzimas com inibidores e outras proteínas. Todos os OGMs são de Classe de risco I. Cepas de *E. coli* não patogênicas DH5 $\alpha$  e BL21DE3. Vetores expressão de proteínas pET28a, pET29, pET-TEV, pGEX e pPROEX e propagação pUC57. Material genético incluído no OGM: Os genes que codificam as proteínas alvo de estudos desse projeto, no genoma humano, já foram identificados e as sequências de nucleotídeos e de aminoácidos estão disponíveis nos bancos de dados GeneBank ([www.ncbi.nlm.nih.gov/gene/](http://www.ncbi.nlm.nih.gov/gene/)) e Uniprot ([www.uniprot.org/](http://www.uniprot.org/)), respectivamente.

O projeto de pesquisa acima descrito, a ser conduzido pela aluna de doutorado Valéria Scorsato, recebeu autorização desta Comissão Interna de Biossegurança.

Presidente em exercício da CIBio-IQ: Carlos H. I. Ramos