PRE-CLINICAL ASSESSMENT OF THE PROTEASOMAL INHIBITOR BORTEZOMIB AS A GENERALIZED THERAPEUTIC APPROACH FOR RECESSIVELY INHERITED DISORDERS

Calvin Jary

This thesis is submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the Master's degree in Cellular and Molecular Medicine

Cellular and Molecular Medicine

Faculty of Medicine

University of Ottawa

Abstract

The number of known monogenic rare diseases (~7000) exceeds the number of effective treatments (~500) by more than an order of magnitude underlining the pressing need for generalizable therapeutic approaches for this class of conditions. In this regard, the majority of recessive and x-linked recessive disorders are caused by missense mutations encoding proteins that frequently have residual function but are rapidly degraded by the 26S proteasome. Bortezomib is a small molecule that inhibits the 26S proteasome and has been approved for use in patients for an unrelated condition; multiple myeloma. Previous work has shown that, for a small number of disorders, bortezomib can inhibit the degradation of the mutant protein, thereby increasing the protein level and activity, holding out the promise of a beneficial therapeutic effect by the repurposing of this agent. We present here a high level western blot based survey of nine recessive disorders to characterize the general effectiveness of such an approach. Thirteen patient fibroblast cell lines comprising 9 different diseases with 19 known mutations were selected on the basis of missense mutations protein expression data when available. The cell lines were incubated with bortezomib (10 nM and 50 nM; 24 hrs) and levels of the mutated protein were quantified by western blot. Unfortunately, no consistent, appreciable increase was observed for any of the conditions tested. The general therapeutic value of re-purposing bortezomib for recessive and x-linked diseases appears limited at best. The few reported cases of bortezomib successfully working in increasing mutated protein levels appear to be the exceptions and not the norm. Moreover successes are more often limited to cell lines carrying a transgene expressing the mutated protein rather than endogenous mutated protein in patient cell lines.

Table of Contents

Preliminary Pages

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of tables

List of Figures

List of abbreviations

BSE	Bovine Spongiform Encephalopathy
DNA	Deoxyribonucleic Acid
ER	Endoplasmic Reticulum
ETC	Electron Transport Chain
FCS	Fetal Calf Serum
G	Gibb's Free Energy
H	Enthalpy
kDa	Kilodaltons
mRNA	Messenger RNA
NHF	Normal Human Fibroblasts
nM	Nanomolar
RIPA	Radioimmunoprecipitation Assay
S	Svedberg
SD	Standard Deviation
SNP	Single Nucleotide Polymorphism
T	Temperature
vCJD	variant Creutzfeldt-Jakob disease
WT	Wild type
µg	Microgram
µL	Microliter
µM	Micromolar

Acknowledgements

I would like to sincerely thank my thesis supervisor Dr. Alex MacKenzie for his unwavering support, both moral and scientific, throughout the struggles and successes of this study, he was truly a joy to work with. Thanks also goes to the entire MacKenzie lab and especially to several members of the lab who provided key insight and without whom I could not have succeeded. Namely Sean O'Reilly, Kevin Mongeon and Dr. Izabella Pena.

I would also like to thank everyone at the Apoptosis Research Centre, especially Lynn Kelly, who make the whole facility run smoothly by placing orders and answering questions, and contributed to a great environment. I am very appreciative of my advisory committee members Dr. Brian Foster and Dr. Michael Geraghty for their guidance during the course of my time at the University of Ottawa.

Finally I would like to thank my mother, for her unwavering support, her endless sacrifices and her desire to see me do what I enjoy.

CHAPTER 1

1.0 Introduction

1.1 Rare Diseases

The latest estimates place the number of rare disease at six to seven thousand, (defined as conditions affecting less than 1 in 2000 people worldwide)(Orphanet 2016a). Almost all genetic diseases are classified as rare diseases, however there are other causes of rare diseases as well, including those of an infectious, auto-immune, or cancerous nature. Most rare diseases, regardless of etiology, are chronic, serious and progressive (Orphanet 2016c). Furthermore, only about 400 of these diseases have approved treatments, thus leaving some 94% untreated (NIH 2016). The genetic rare disease can further be classified by their mode of inheritance, with 28% being autosomal recessive and a further 7.0% being X-linked. These groups are of potential therapeutic interest given that treatments that might upregulate the disease gene or otherwise increase the amount of the existing protein could potentially have a positive impact on the disease, since lack of protein or the lack of protein function is responsible for the disease phenotype. The other modes of inheritance such as autosomal dominant at 26% and multigenic/multifactorial at 13.4% would either be worsened by any successful protein upregulation or be largely unaffected due to too many genes and proteins being involved in the disease phenotype (Orphanet 2015).

With respect to the types of genetic mutations that occur, the overwhelming majority of genetic diseases are caused by single base pair substitutions, also known as missense mutations. Specifically, a landmark study of 18,200 different disease causing mutations in 895 human genes revealed that most mutations that caused disease were single base pair substitutions at 71%, with small deletions and insertions of under 20 base pairs being the next most prevalent at 22.5% and

significant gene deletions and insertions of over 20 base pairs being uncommon at 6.5% and comprising the remainder of the types of mutations (Krawczak et al. 2000).

There is thus a significant demand for a large number of therapies, as thousands of rare diseases have no treatment, however the total number of patients, and therefore the potential market, for any one rare disorder is relatively small. Thus a single therapeutic technology that could work for numerous rare diseases would be ideal, the ultimate goal being perhaps gene therapy or cell therapy tailored to the patient. In the meantime repurposing drugs for select rare diseases remains a lower-cost and feasible option.

Another potential generalization to be made with recessive genetic disorders, is that they are loss-of-function, that is to say; the patient has inherited two altered copies of a gene resulting in either a compound heterozygote or homozygous missense mutation. As described earlier, this is statistically likely to be a single base pair substitution, and also more likely to occur not in the relatively few amino acid residues that comprise the functional domain of the protein, but rather in the greater number that comprise the structural bulk of the protein. Furthermore because misfolded proteins are targeted for degradation and have a reduced biological half-life, it is possible the disease phenotype arises not solely from the mutated protein's reduced function, but also from its active degradation by the proteasome. For example in the case of in CAPN3, a type of muscular dystrophy, about 20-30% of patients have normal levels of the CAPN3 protein, implying mutations in the functional domain and lack of receptivity to any treatment that would increase the amount of mutant protein (Fanin & Angelini 2015). The remaining 70-80% should be composed of missense mutations that should lower protein levels and be receptive to bortezomib treatment, as well as about 6.5% that are likely large additions or deletions which could trigger premature stop codons and result in nonsense mediated decay.

1.2 Protein Folding and Degradation

1.2.1 Protein Synthesis and Introduction to Folding

When a protein is synthesized, secondary structure formation begins almost immediately as predominantly α-helices and β-sheets, with tertiary structure occurring later, being more dynamic, and at times requiring the binding of chaperone proteins to make their correct folding more thermodynamically favourable, ideally with an absence of gene sequence alterations, temperature shock and oxidative stress (Hartl & Hayer-Hartl 2009). Perhaps counterintuitively, the overall energy and forces holding a protein in its final folded state are quite weak, this is due to most proteins being required to maintain significant conformational flexibility in order to function. Furthermore about 20-30% of proteins found in mammalian cells are devoid of any inherent tertiary structure during their synthesis, and entirely depend upon chaperone proteins (Dunker et al. 2008). The wild type conformation of a protein is often a free-energy minimum, which is how the wild type conformation is arrived at in the first place. However there are global minima as well, and these are generally protein aggregates and plaques which are the basis for disease, and demonstrated as the lowest energy structures shown in (Figure 2) (Bross et al. 1999). Should the process of tertiary folding go awry, the protein can be broken down into short amino acid residue chains and synthesis can begin again (Bross et al. 1999). This is accomplished for approximately 80-90% of all proteins by the 26S proteasome, the other 10-20% of proteins, particularly proteins of the cell membrane, are degraded by the lysosome (Sorokin et al. 2009). These two fates are collectively referred to as protein quality control. Dysfunction in

protein folding and the degradation and recycling of the misfolded proteins gives rise to a wide range of diseases categorized as proteopathies.

There are three separate types of misfolded proteins that evade the initial protein quality control machinery: The first type are proteins with a stable, misfolded conformation which results in dominant negative effects in the cell, where the protein not only does not perform its function but causes harm to the cell, for example epidermolysis bullosa simplex (Bross & Gregersen 2003). The second type are misfolded proteins which may form diffuse aggregates also referred to as amyloids, in which newly synthesized proteins are added to a protein mass which forms a plaque of sorts, Alzheimer's is a prime example. The third and final type is the case for autosomal recessive orders, a mutated protein that is not degraded simply persists in the cell unable to perform its role, and as such should both alleles code for mutated proteins, a disease by functional deficiency will occur, for example cystic fibrosis (Figure 1)(Bross & Gregersen 2003).

The model of simplifying all protein misfolding fates into one of these three types is quite versatile, and can even explain more exotic instances of mutant protein dysfunction such as BSE and its human equivalent vCJD. This disease is an instance where a stable misfolded protein with dominant negative effects converts wild type proteins to its mutant configuration, and together they form amyloids which damage the brain, and are thus simply a combination of the first and second type of mutant protein activity. At each of these three misfolded conformation types, there is some steady state levels of protein quality control activity which is degrading the misfolded proteins and is turning them back into short polypeptides for novel protein synthesis.

Impaired folding is of great interest when seeking to treat diseases of functional deficiency generally and autosomal recessive disorders specifically. The vast majority of

functional deficiency due to protein misfolding occurs due to accidental misfolding of a wild type polypeptide sequence, and therefore the protein quality control machinery is the primary mechanism in disposing of such a misfolded protein so it does not stress the endoplasmic reticulum nor run the risk of forming amyloids. However should mutated misfolded protein retain some function and be appropriately targeted within the cell, degradation by protein quality control will only exacerbate the phenotype and have a deleterious effect on the patient. In such a circumstance, inhibiting the proteasomal degradation functionally competent proteins may be of therapeutic value.

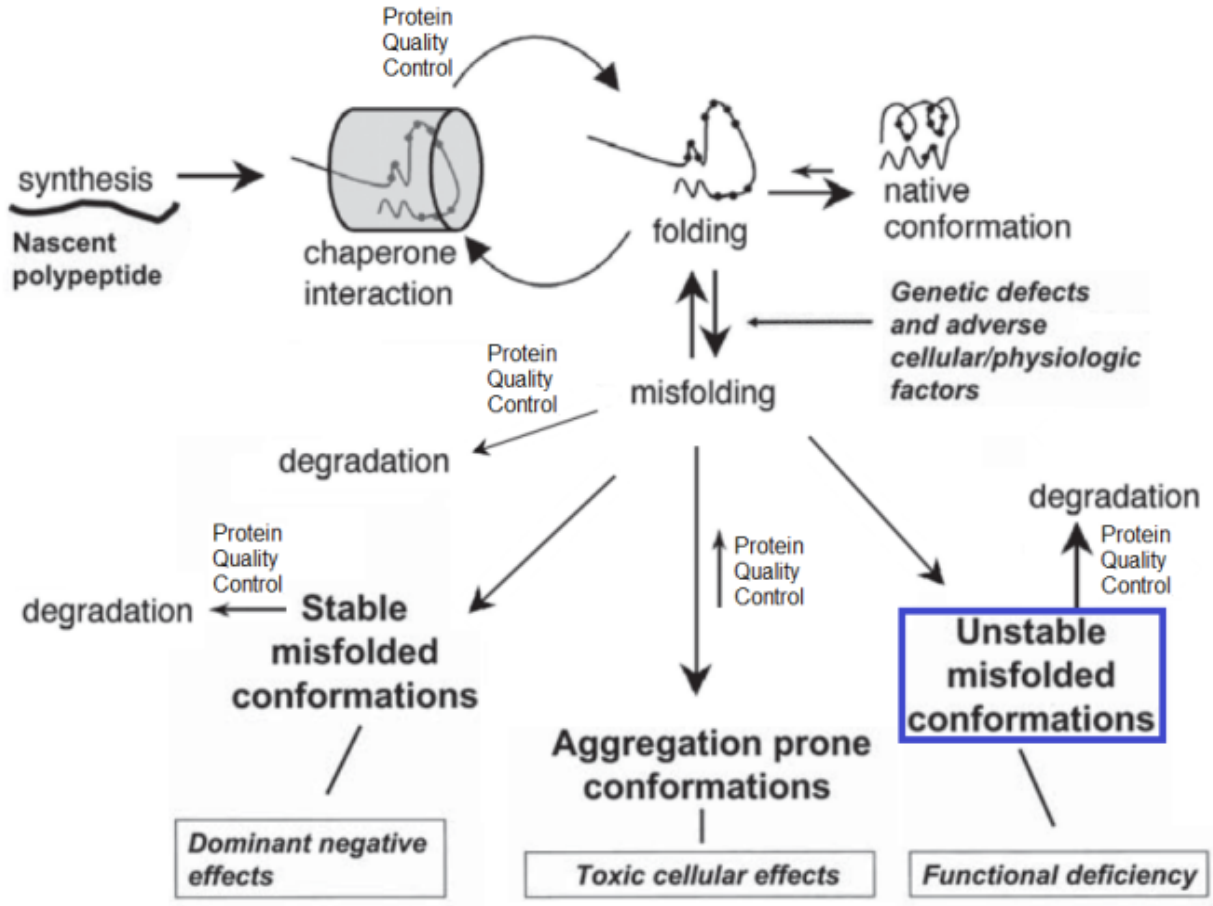


Figure 1. **An extensive diagram of polypeptide fate**. Recessive disorders are highlighted by the blue rectangle. Bortezomib's role is to inhibit degradation of unstable misfolded conformations by the quality control machinery. This figure has been extensively modified from its original source (Bross & Gregersen 2003).

1.2.2 Mechanisms and Biochemistry of Protein Folding

While proteins are correctly characterized as being the most versatile and structurally complex of the four classes of macromolecules they are often incorrectly characterized as being structurally rigid and having a native state that is a thermodynamic inevitability arrived at independently within microseconds. The reality is proteins must remain conformationally flexible to fulfill their catalytic or mechanical roles and as a result maintain only marginal thermodynamic stability even when correctly folded. Moreover, while small proteins can fold within microseconds, larger proteins not only can take minutes or hours to fold. Even with the presence of chaperone proteins, high cytosolic protein concentrations (300-400 g/L) predispose non-native conformations and protein aggregation, particularly of long hydrophobic domains. Correct folding is far from a thermodynamic inevitability as well as folding interactions are highly complex, heterogeneous and rely on the cumulative effects of many weak, non-covalent interactions including hydrophobic interactions which help bury non-polar amino acid residues into the interior of the protein, while avoiding the deleterious hydrophobic interactions which occur on the surface of proteins and can lead to irreversible amorphous aggregates and amyloid fibrils predisposing cells to diseases such as Alzheimer's. Furthermore there are an astronomical number of ways a protein can remain "stuck" as a folding intermediate or in some partially folded state, requiring either additional time, the action of chaperone proteins or ultimately protein degradation and *de novo* synthesis by the cell (Figure 2). Specifically, approximately 20-30% of eukaryotic proteins do not naturally assume a tertiary structure and require binding partners to adopt such a structure, namely chaperone proteins (Dunker et al. 2008).

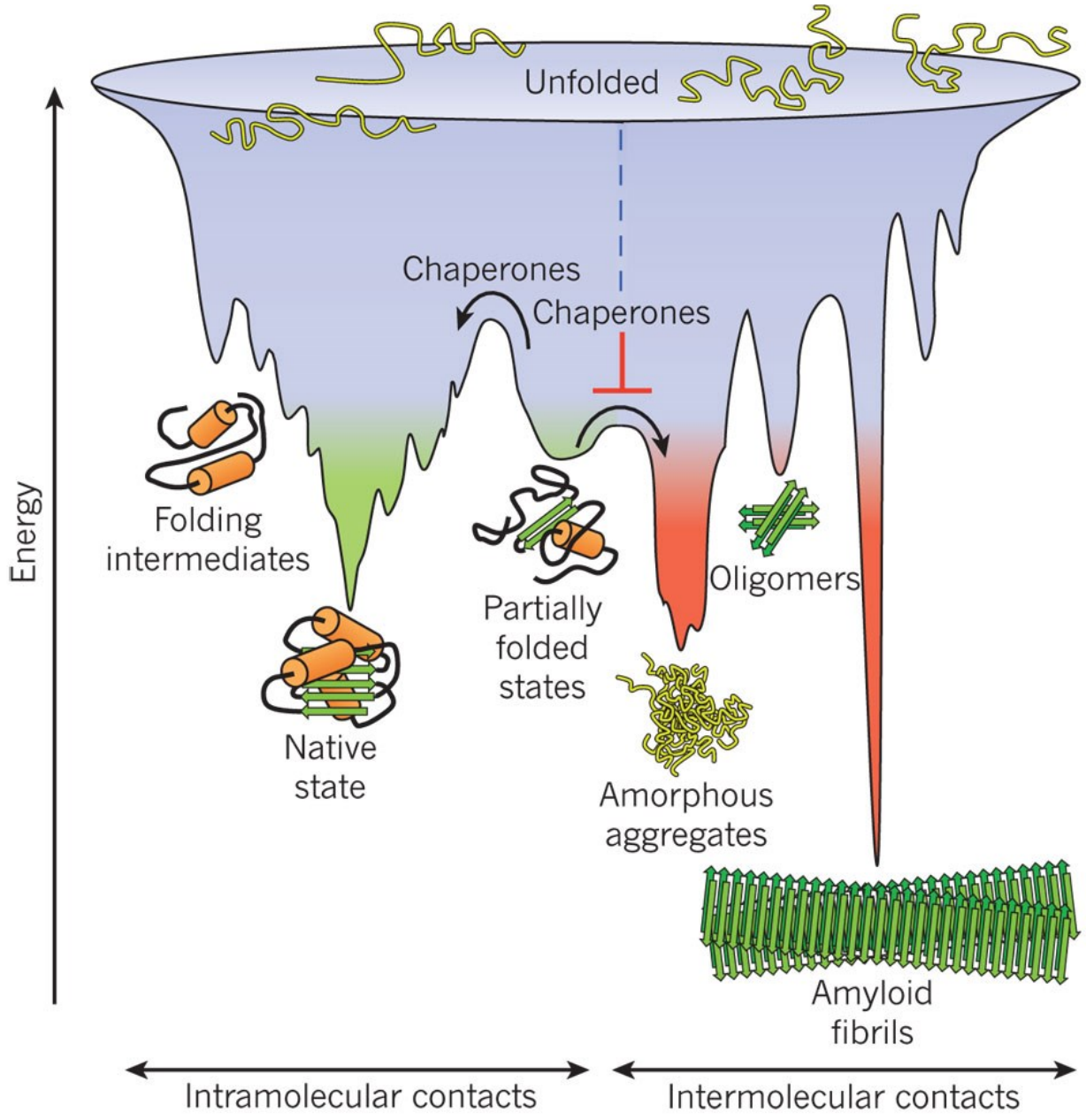


Figure 2. The free-energy surface that a protein explores as it folds. The green coloured native state is achieved by forming chaperone-assisted intramolecular contacts and by avoiding chaperone-inhibited intermolecular contacts that lead to aggregation and disease (Hartl, F., Bracher, Andreas., Hayer-Hartl 2011).

1.2.3 Thermodynamics of Protein Misfolding

When proteins are transcribed from RNA, they begin forming secondary structures; chiefly alpha helices and beta sheets immediately and as the transcription process is happening. However the formation of their tertiary structure takes much longer and requires the peptide synthesis to be complete in order to occur (Hartl, F., Bracher, Andreas., Hayer-Hartl 2011). Significant disruption of proper protein folding due to missense mutations generally happens in one of two ways, either a hydrophobic amino acid residue in the core of the protein is changed to a hydrophilic one, which is thermodynamically destabilizing potentially exposing parts of the core to the hydrophilic surface environment, or alternatively, a hydrophilic surface residue is changed to a hydrophobic one, causing thermodynamic instability on the surface of the protein. There are of course a multitude of less drastic missense mutant replacements which are also less drastic in their disruption of protein folding. Thermodynamic instability and destabilization naturally translates into protein misfolding which in turn leads to degradation. However a way of quantifying this effect is necessary as this is not a binary effect and not all mutations are equally destabilizing.

The best physical quantification of protein misfolding and how mutations might effect thermodynamic severity is with the Gibb's free energy equation: $\Delta G = \Delta H - T\Delta S$. This equation allows for the accurate measurement of the change in a protein's stability as a result of a new condition. The root equation $G = H - TS$ accounting for the total energy value of a protein cannot be used, because total enthalpy (H), can never be accurately measured as there is no universal reference frame. When comparing the stability of a missense mutation versus its wild type counterpart, the change in ΔG between the two proteins is appropriate to use, which is noted as $\Delta\Delta G$ (Dill 1990).

To quantify the degree of destabilization missense mutations cause we can rely on ΔΔG and experimental studies performed with large numbers of proteins. An impressive study was done across 16 different globular, single domain, monomeric proteins of varying sizes (from 50 to 330 residues) which were subjected to 1285 experimental mutations (Tokuriki et al. 2007). The resulting ΔΔG energy changes and how common they were, smoothly fit a Gaussian distribution (Figure 3). Qualitatively, about 8-29% of missense mutations were found to be stabilizing and 4-45% of missense mutations were found to be destabilizing, with the remaining missense mutations being neutral. This large range is indicative of how unique the characteristics of each protein are. Indeed, no universal rule could be found to explain these large differences between proteins, let alone predict such distributions (Tokuriki et al. 2007).

Despite this large distribution in ΔΔG values, there is a clear difference between mutations of the surface hydrophilic amino acid residues which have a mean of about ~0.6 kcal/mol and are mildly destabilizing, and core hydrophobic residues which have a mean of about ~1.4 kcal/mol and are significantly destabilizing. For context, these globular proteins have a total tertiary structure thermodynamic stability in the range of only -5 to -15 kcal/mol, and a single intraprotein hydrogen bond is only 1-2 kcal/mol (Lodish H, Berk A 2000). This low thermodynamic tertiary structure stability results in protein flexibility necessary for appropriate function. A corollary of this concept is that a typical missense mutation disrupting a single hydrogen bond, considering a globular protein's tertiary structure is maintained by only a handful of hydrogen bonds worth of energy, can easily result in a misfolded protein (Dill 1990; Tokuriki et al. 2007).

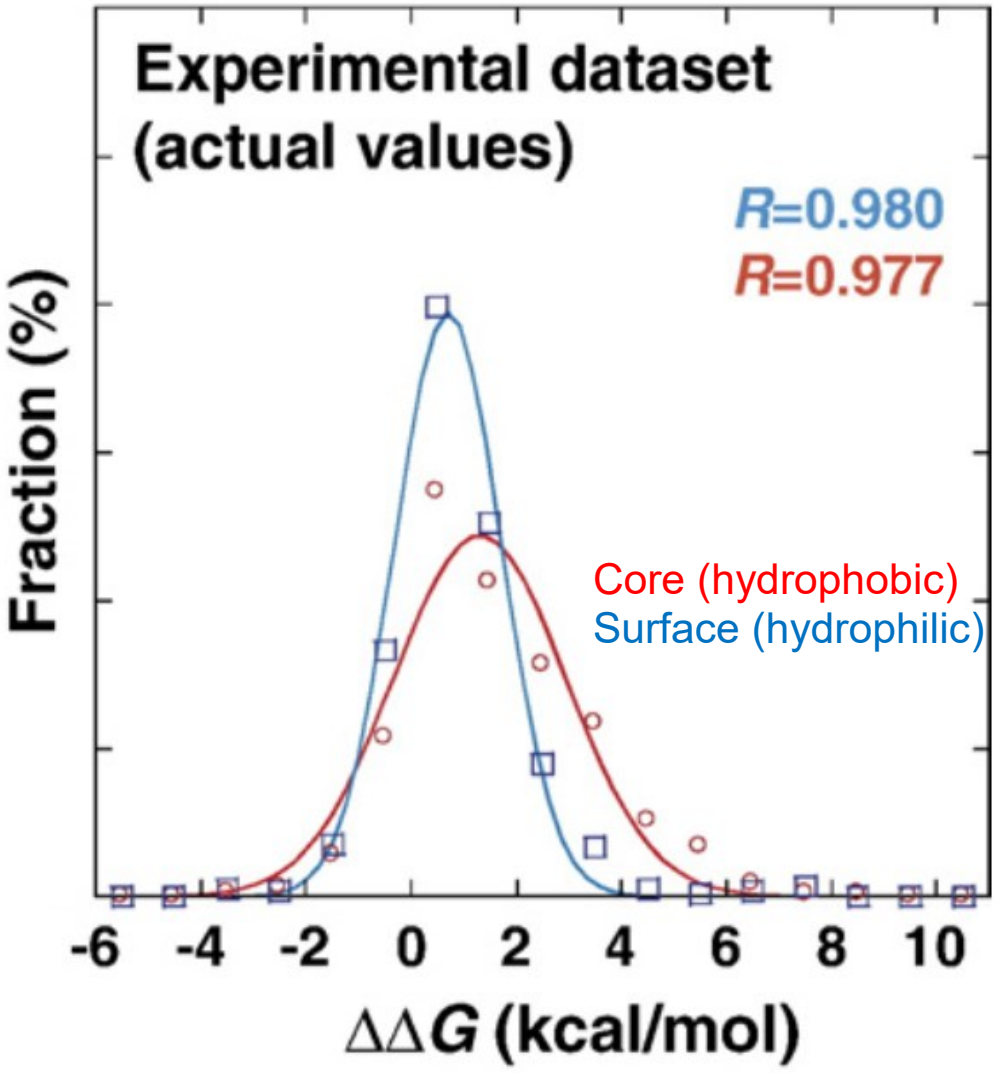


Figure 3. Individual ΔΔG distributions of core and surface residues. The distributions were fitted to a single Gaussian function. Mutations of hydrophobic amino acids in the core of the protein generally had a greater destabilizing effect than mutations of the surface hydrophilic amino acids. The R value represents the amount of correlation to the fit of the shown Gaussian function (Tokuriki et al. 2007).

1.2.4 Protein Degradation and the 26S Proteasome

Of the 80-90% of proteins that are degraded by the proteasome, the mechanics are simple conceptually, but are more complex in detail. The exact mechanism as to why a protein is tagged for degradation is not well known, but there are multiple compelling observations. There is the N-end rule, that proteins with a N-terminal consisting of methionine, serine, alanine, threonine, valine or glycine generally have half-lives greater than 20 hours, whereas those with a N-terminal of phenylalanine, leucine, aspartic acid, lysine or arginine generally have half-lives of 3 minutes or less. This is one simple strategy a cell can use to be able to create and regulate both long-lasting housekeeping proteins and more dynamic proteins whose levels can rapidly be modulated (Varshavsky 2011).

More relevant for missense mutations and recessive disorders however is the PEST domain, which consists of approximately 8 contiguous residues which are rich in Pro, Glu, Ser and Thr, and with the presence of this domain resulting in a protein that is far more likely to tagged for degradation. For example the protein Gcn4p goes from a 5 minute to a 50 minute half-life when its PEST sequence is removed. The machinery that identifies a protein as misfolded appears to be triggered by the presence of hydrophobic residues on the surface of a given protein. (Rechsteiner & Rogers 1996). For the most part however, the mechanism underlying "why" a given protein is tagged for degradation remains largely unknown.

The "how" of degradation is well understood, starting with multiple 8.5 kDa regulatory proteins called ubiquitins being added to a particular amino acid residue of the target protein. The ubiquitin proteins are first activated by ubiquitin-activating enzymes (E1s) then conjugated by ubiquitin-conjugating enzyme (E2s) and then finally conjugated to the particular residue on

the target protein by ubiquitin ligase (E3). When this process is performed as a monoubiquitination, whereby a single ubiquitin is added to any number of potential sites on the target protein, the process has functions in membrane trafficking, endocytosis, inflammation, translation and DNA repair. (Mcdowell & Philpott 2013)

When protein degradation is required, polyubiquitination must take place. This occurs when multiple ubiqituins are added in a linear chain to a particular lysine residue on the protein in question; each successive ubiquitin bonded to specifically the lysine 48 of the preceding member. Although ubiquitin contains seven lysine residues, lysine 48 is the sole anchor point which leads to proteasomal activity. Once at least four ubiquitins have been added in this manner, proteolysis is signaled (Varshavsky 2011).

Once tagged for degradation the 26S proteasome, a cylindrical structure, consisting of a 20S (for Svedberg) core particle and the 19S regulatory particle become activated. The regulatory particle allows only polyubiquitinated proteins to enter the 20S core. The core is quite narrow and incoming proteins are at least partially unfolded and linearized. As they proceed inside the core proteolysis occurs by a nucleophilic attack on threonine residues within the protein, causing breakage of the peptide bond and short polypeptides 4 to 25 residues in length, which may be reused for novel protein synthesis (Figure 4). The detailed mechanism for all of this complex machinery is known and involves dozens of individual proteins and subsystems working in tandem, and has been described in impressive detail (Matyskiela et al. 2013).

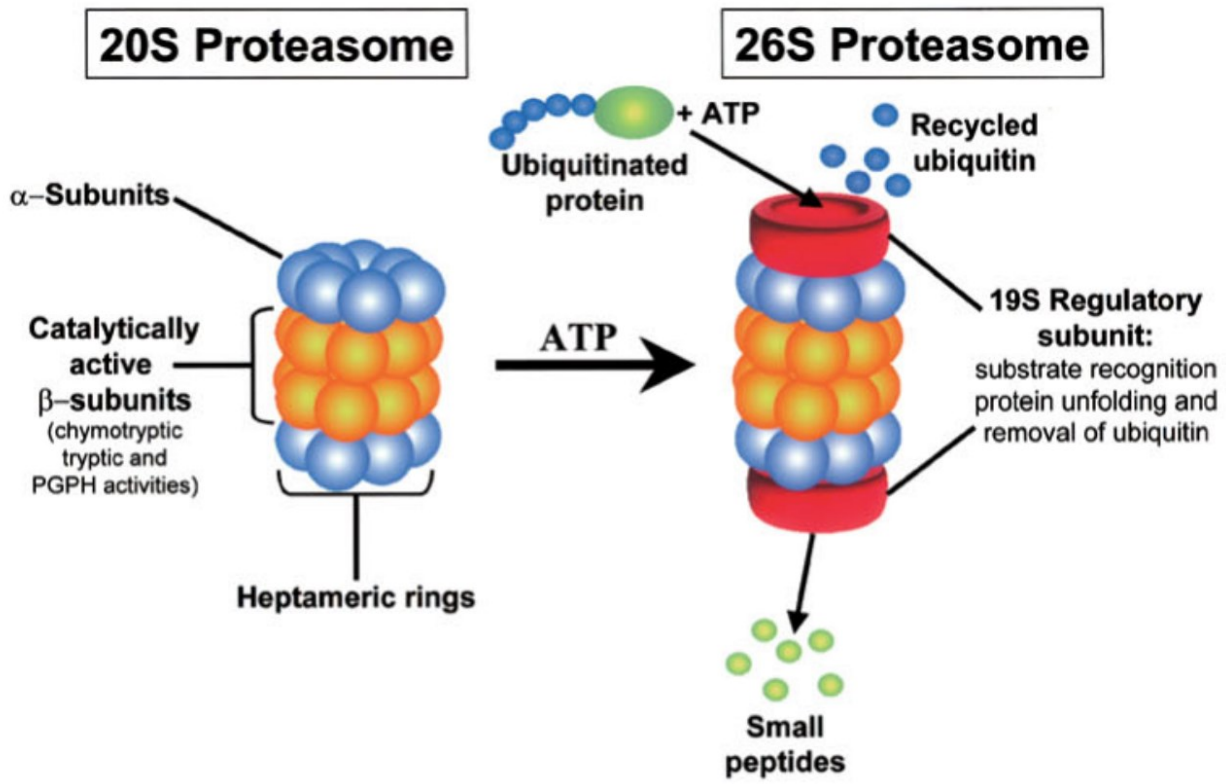


Figure 4. Diagram of the functioning of the 26S proteasome. The main 20S Proteasome subunit (left) is the catalytically active subunit that is the inhibitory target of bortezomib. The 20S core subunit and the 19S regulatory subunits combine to form the 26S proteasome, which cleaves proteins with polyubiquitinated tails in an ATP-dependant manner into small peptides and monomeric ubiquitin tags (Almond & Cohen 2002)**.**

1.3 Bortezomib (VELCADE®)

1.3.1 Introduction to Bortezomib

Bortezomib is a boronic acid dipeptide in classification. It is a highly selective, reversible inhibitor of the 20S core subunit of the 26S proteasome, with a K_i of only 0.6 nM (Janssen 2015; SelleckChemicals 2013). That is to say the binding affinity of bortezomib is such that only 0.6 nM in the cell will inhibit half of all 26S proteasome function. Approximately 83% of the drug binds to plasma proteins, this percentage remains constant regardless of drug concentration. The molecule penetrates the cell membrane readily and reduces the proteolysis of even long-lived proteins by inhibiting the 26S proteasome (Adams et al. 1999). The drug has been approved by the FDA and Health Canada for the treatment of relapsed multiple myeloma, patients that have failed more conventional therapies. Due to the drug's success for relapsed patients it is now a first line therapy for newly diagnosed multiple myeloma patients as of 2009 (Janssen 2015).

The drug is the most practical candidate for interfering with the ubiquitin-proteasome pathway for recessive disorders given its existing approval for human use by all developed countries, though it is not the only proteasome inhibitor. Other 26S proteasome inhibitors also include research-only compounds such as MG-132, MG-115, lactacystin, epoxomicin and PS-341. While some of these compounds could be taken orally allowing for much more convenient administration than bortezomib's requirement for intravenous or subcutaneous injection, the compounds are not currently approved for human use.

It is believed that bortezomib's mechanism of action for multiple myeloma is a global inhibition of protein degradation, resulting in key cell cycle inhibitory proteins persisting longer in the cancer cell and performing their oncolytic function. This includes: p21, p27, p53 as well as additional pro apoptotic proteins such as Bid, Bax, caveolin-1, proapoptotic c-Jun-NH2 terminal kinase and inhibitor κB-α, as well as increasing the overall ER stress response as misfolded and hastily-made proteins persist longer in the ER, which reduces the rate of cell division and increases the chance of cell death (Janssen 2015).

When used as an antineoplastic, a high dose of the drug is given (1.3 mg/m²) which is usually administered intravenously at a concentration of 1 mg/mL, but can also be administered subcutaneously at a concentration of 2.5 mg/m. This dosing results in an approximate, measured, steady-state serum concentration of about 270 nM. This treatment regimen is given across three 3-week periods where the drug is administered on days 1, 4, 8 and 11. For comparison, i*n vitro* testing has shown a 50% growth inhibition across 60 various human tumours with a treatment concentration of 7 nM (Janssen 2015; SelleckChemicals 2013).

Due to its powerful binding affinity of 0.6 nM K_i, when the drug has been repurposed in the literature for the indication of recessive disease treatment, concentrations as low as 10 nM have been found to be effective, this is a mere 3.7% of the multiple myeloma treatment dose or a concentration reduction of a factor of 27. Cell viability assays have been performed with no increase in cell death up to the maximum timepoint of 96 hours at 10 nM (Shimada et al. 2011; Azakir et al. 2012). Bortezomib is metabolized first by slow oxidative deboronation in the presence of water, which inactivates the function of the drug, with 90% of the boron being eliminated renally (Bolognese et al. 2009). Next a large number of cytochrome P450 enzymes

including CYP3A4, CYP2C19, CYP1A2, CYP2D6 and CYP2C9 which results in nine separate metabolities which are also eliminated renally (Pekol et al. 2005).

1.3.2 Bortezomib Impact on Recessive Disease: State of the Art

Some recent publications report that bortezomib's induction of mutant protein levels by significant amounts in patient cells is able to alleviate the disease phenotype (Azakir et al. 2012). While this success is across a diverse spectrum of recessive disorders representing a significant range of protein sizes and subcellular localizations; the number of mutations for a given protein that are shown to be receptive to bortezomib treatment is quite limited, and in most publications is only a single missense mutation. These examples include Pompe disease (Glycogen Storage Disease Type II) a disease of acid alpha glucosidase (the GAA gene) which is responsible for the degradation of glycogen to glucose in lysosomes. The c.546G>T mutation was successfully treated with a 10 nM dose of bortezomib over 24 hours in patient fibroblasts. This resulted in a 4-fold increase in the amount of the mutant GAA protein, from 7.3% wild type levels up to 29.8%. Furthermore there was greater co-localization of the produced GAA within lysosomes, the target organelle of GAA (Shimada et al. 2011).

Dysferlinopathy (Limb Girdle Muscular Dystrophy Type 2B) is a disease of dysferlin (the DYSF gene) which is responsible for cell membrane regeneration and repair. The Arg555Trp mutation was successfully treated with a 50 nM dose of bortezomib over 24 hours in 180/06 myoblasts. This resulted in a massive 24-fold increase in the amount of the mutant dysferlin protein, as well as a 4-fold increase in mRNA. Interestingly the disease phenotype was also treated as 10 nM of bortezomib treatment for 24 hours in the 180/06 cells allowed the mutant dysferlin to significantly repair the plasma membrane within 5 minutes of a laser

inducing holes in the membrane, whereas without bortezomib, repairs had not even begun after 5 minutes (Azakir et al. 2012)

Niemann–Pick disease type C1 is a disease of NPC Intracellular Cholesterol Transporter 1 (the NPC1 gene) which is an intracellular cholesterol mediation protein. The Gln775Pro, Ala1035Val, Ile1061Thr, Thr1066Asn, Leu1106Pro and Arg1186His were all successfully treated with a 30 nM dose of bortezomib over 24 hours in NPC patient fibroblasts. These treatments resulted in between a 1.3 to 5.3-fold increase in the amount of mutant NPC protein. In the case of Gln775Pro and Ala1035Val in particular, mutant protein levels went from ~30% to nearly 100% (wild type levels) effectively treating the disorder, however functional assay tests were not performed (Macías-Vidal et al. 2014).

In addition to these three publications that have shown bortezomib effective in patient fibroblasts and myoblasts, there are publications which perform transgenic experiments, whereby the gene for the mutant protein, and the gene for the wild type protein for comparison, are transfected into various cell lines, not all of them human, and the bortezomib efficacy for mutant protein upregulation is once again quantified. These include Gunther's disease (Congenital erythropoietic porphyria) is a disease of uroporphyrinogen III synthase (the UROS gene) which is a protein that catalyses the fourth step of porphyrin biosynthesis in the heme biosynthetic pathway. K562 cells, which are human immortalized myelogenous leukemia cells, were stably transfected with Cys73Arg and Pro248Gln mutants in the form of plasmids with eGFP fused to the C terminus. Upon treatment with 100 nM of bortezomib for 16 hours, there was a 53-fold increase in the former and a 30-fold increase in the latter mutation (with the included eGFP on the C terminus), relative to the wild type protein which was also transfected and also contained the eGFP (Blouin et al. 2013).

Friedreich Ataxia is a disease of frataxin (the FXN gene) which is a protein that regulates mitochondrial iron transport and respiration. When His-frataxin was transfected into Calu-6 cells (cells of an anaplastic carcinoma of lung origin, chosen for their high transfection efficiency and elevated frataxin expression), treatment with 10 μM of bortezomib for 18 hours caused a 2.6-fold increase in the amount of the His-frataxin (Lavecchia et al. 2013).

Sanfilippo syndrome (Mucopolysaccharidosis type IIIA) is a disease of N-sulfoglucosamine sulfohydrolase or sulfamidase (the SGSH gene) which is one of several enzymes involved in the lysosomal degradation of heparin sulfate. When the mutant Ser298Pro and wild type SGSH gene were transfected into BHK (baby hamster kidney) cells, 50 μM of ALLN and also 10 μM of lactacystin (both proteasome inhibits like bortezomib) over 24 hours were able to increase the amount of the overexpressed Ser298Pro protein levels from approximately 3% wild type levels to what visually appears to be wild type levels (Muschol et al. 2011).

Cereblon deficiency (Autosomal Recessive nonsyndromic intellectual disability) is a disease of cereblon (the CRBN gene) which is a substrate receptor for the Cullin-RING E3 ligase complex. When the mutant Arg419X and wild type gene for CRBN were transfected into HEK293T (human embryonic kidney) cells and treated with bortezomib at 1 μM for 24 hours, there was an 8-fold induction in the mutant protein which almost matched the wild type levels of CRBN transfection (Xu et al. 2013).

1.3.3 Risks of Bortezomib for Recessive Disorders

It is likely that using 10 nM Bortezomib (rather than the oncological 270 nM) is a safer concentration, this improvement in safety via a large decrease in concentration may be offset in treatment duration, from the oncological three 3-week treatments to a lifelong regimen.

Additional concerns for lifelong administration for recessive disorders, is that the partial inhibition of the proteasome's key role in protein quality control would reduce the cell's efficacy in combating aggregate prone protein conformations, and to a lesser extent dominant negative effects caused by stable misfolded conformations. Since the drug does not penetrate the blood-brain barrier, there should be no increased risk of neurodegenerative disorders caused by protein aggregates, such as Alzheimer's, Parkinson's and Huntington's, as bortezomib might increase their risk should it be able to penetrate the barrier (EMEA 2004).

However there may be concerns of protein aggregate propensity in other tissues with bortezomib treatment. Specifically muscular tissues and their increased susceptibility of protein aggregate myopathies (PAMs), which are numerous and include desminopathies, actinopathies, myosinopathies and others. Fortunately the literature suggests that when protein aggregate myopathies arise, they do so in mutant muscle proteins, and partially inhibiting the protein degradation mechanisms would not be sufficient to trigger these disease phenotypes in patients which produce wild type muscle proteins. This is due to the presence of hydrophobic amino acid residues on mutant protein surfaces, causing the protein-protein interactions that lead to insolubility and aggregation (Bross et al. 1999; Sharma & Goebel 2005). This is however a theoretical argument, and long-term safety of bortezomib at any dose has not been demonstrated in the literature. Finally bortezomib has no mechanistic basis to correct any potential improper shuttling of a mutant protein, additionally the protein quality control machinery could be holding

the mutant protein in the endoplasmic reticulum. Furthermore, the multiple ubiquitin protein tags attached to the mutant protein, originally targeting it for degradation, may impact correct shuttling, localization, and functioning of the mutant protein. However there is at least some evidence that most missense mutations are properly shuttled (Bross et al. 1999), and additional evidence that in some cases, perhaps cherry-picked, protease inhibitor treatment significantly improves the shuttling prospects of mutations that indeed experience shuttling dysfunction (Shimada et al. 2011; Hirayama et al. 2008).

1.4 Study Rationale

Ideally, a significant fraction of recessive disorders caused by compound heterozygous or homozygous missense mutations will be treatable with Bortezomib, as there is a greater probability of a missense mutation not being in the functional domain of a protein and causing some protein misfolding and ensuing degradation, than the mutation occurring on a functional domain within the protein rendering it inactive regardless of its fate by way of proteasomal degradation. While there should be some safety improvement with a 27-fold drop in treatment dose from bortezomib's antineoplastic approved dose to this recessive disorder indication, there is also increased risk as the treatment duration would be expanded from 4-6 months to a potentially lifelong duration. By increasing the quantity of mutant protein in patients by several fold, as demonstrated in the literature for specific mutations, we might effectively alleviate the disease phenotype and provide treatment to patients across a diverse range of recessive disorders which might otherwise go untreated.

1.5 Hypothesis and Objectives

The hypothesis is that missense mutated but nonetheless functional proteins underlying recessive disorders will be effectively increased by inhibiting their proteasomal inhibition, as has been showing in the literature in a few specific cases (Shimada et al. 2011; Azakir et al. 2012; Macías-Vidal et al. 2014)..

Therefore this study has two mutually exclusive objectives:

i. Attempt to treat numerous promising missense mutations of qualifying recessive disorders with bortezomib and measure any success by significantly more protein present in the treated cells by Western blot. If receptive mutations are found, continue with examinations of protein subcellular localization and functional assay to demonstrate alleviation of disease phenotype.

ii. Should no missense mutation that is impacted by bortezomib treatment be found, attempt to explain and cast light on the few successful cases of bortezomib in the literature not being representative of the broad recessive disorder landscape.

CHAPTER 2

2.0 Methods

2.1 Locating and Filtering Rare Disease Candidates for Potential Bortezomib Therapy

A key source of information on rare diseases is a rather exhaustive list published by Orphanet, a version of which can now be found arranged by decreasing disease incidence. This list includes approximately 4000 rare diseases in alphabetical order (Orphanet 2016c). This encompasses most, but not all, of the approximately six to seven thousand rare diseases that exist. Further refining of disease candidates was necessary and proceeded according to the following hierarchy:

1. A prevalence of at least 1 in 100 000:

When sorted by prevalence, Orphanet provides a list of approximately 600 rare diseases which are listed that occur with a frequency of at least 1 in 100 000 in the population (Orphanet 2016b). This prevalence was chosen due to a greater number of patients which stand to benefit is a treatment is found, and a greater ease in acquiring cell lines to experiment with.

2. A disease type which is recessive (autosomal or X-linked):

Recessive diseases shall potentially benefit from a proteasomal inhibition conferred increase in protein while diseases which are dominant, sporadic, multi-factorial or unknown in their genetics make poor candidates.

3. No or minimal other treatment:

If a treatment exists for a particular disease that is superior to bortezomib treatment, certainly that disease should be skipped; unfortunately this is rarely the case. Although diseases with expensive therapies which are effective but out of the reach of many patients might also be candidates for bortezomib.

4. Significant protein expression in human fibroblasts

Fibroblasts are the cell type of choice for this study due to their availability from cell repositories and ease of use. Expression level in fibroblasts for a given gene was provided by SAGE (Serial Analysis of Gene Expression) (GeneCards 2016).

2.2 Cell Culture

All cell lines were grown on 15 cm plastic plates (Cell Star and Bio One) in 20 mL of Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 1% (v/v) L-glutamine, 10% (v/v) fetal calf serum (FCS) (Sigma, F1051) and also with 100 units/mL of penicillin/streptomycin (HyClone, SV30010). Cells were grown at 37°C in a 5% CO2 atmosphere. Culture media was changed 2-3 times per week and cells were passaged 1:2 or 1:3 to maintain an approximate 30 to 80% confluency. The passaging was performed by incubation with 5 mL of pure trypsin for 3 to 5 minutes with subsequent addition of 15 mL of the aforementioned growth media. All cell lines used in this study were human fibroblasts obtained from Coriell Cell Repository (Camden, New Jersey).

Gene name	Disease name	Protein name	Mutation	Cell Origin
ADA	severe combined immunodeficiency disease (SCID)	Adenosine deaminase	R101W & R211H	GM02605
ADA	severe combined immunodeficiency disease (SCID)	Adenosine deaminase	L107P & R211C	GM04395
ADA	severe combined immunodeficiency disease (SCID)	Adenosine deaminase	A329V ex4del	GM02824
ADA	severe combined immunodeficiency disease (SCID)	Adenosine deaminase	R101Q & IVS1(+1GT>CT	GM02445
ALDH3A2	Sjogren-Larsson syndrome	Aldehyde dehydrogenase 3 family, member A2	434X & 457X	GM10640
ALDH3A2	Sjogren-Larsson syndrome	Aldehyde dehydrogenase 3 family, member A2	2x A314G-Ala-Lys-Ser-Thr-Val-Gly-P315A	GM10641
ASS1	citrullinemia	argininosuccinate synthase 1	unknown	GM01058
CAPN3	limb-girdle muscular dystrophies type 2A	Calpain 3	unknown	GM03932
CHM	choroideremia	Rab escort protein 1	unknown	GM06940
DHCR7	Smith-Lemli-Opitz syndrome (SLOS)	7-dehydrocholesterol reductase	45,XY,t(13;14)(13qter>cen>14qter	GM03044
FKRP	muscular dystrophy-dystroglycanopathy	Fukutin-Related Protein	2x L276I	GM23868
GAA	Pompe's disease	acid alpha-glucosidase,	c.-45T>G & c510T>C	Local Patient
SGSH	Sanfilippo syndrome type A	Heparan Sulfate Sulfatase	S298P & R74C	(Muschol et al. 2011)

Table 1. Summary of all genes, known mutations and cell origins for the study.

2.3 Bortezomib Treatment

Cells were plated into 6-well plastic plates (Cell Star and Bio One) at a confluency of approximately 70%. 24 hours later, the media was replaced with 2 mL of fresh media, and 2 µL of pure DMSO, 2 µL of 10 µM bortezomib or 2 µL of 50 µM bortezomib for a final treatment dose of 0 nM (control), 10 nM and 50 nM of bortezomib respectively. The cells were then incubated under aforementioned growth conditions for 24 hours and harvested.

2.4 Protein Extraction

Growth media was removed from cells, the cells were then washed twice with 1X PBS and lysed using 120 µL of RIPA buffer supplemented with protease inhibitors (cOmplete Tablets from Roche). The cells were then incubated for 30 minutes at 4°C to allow for cell lysis and then were scrapped via a cell scrapper and pipetted into 1.5 mL microcentrifuge tubes. Samples were then sonicated for two rounds of 15 seconds each while being kept on ice (Sonics vibra cell). Samples were then centrifuged at 4°C for 40 minutes at 13 000g. The supernatant was pipetted off and placed in a fresh 1.5 mL microcentrifuge tube and stored at -80°C.

2.5 Western Blotting by Stain-Free

Protein concentrations were determined by Bradford assay using a Bio-Rad protein assay kit (Bio-Rad #500-0006). SDS-PAGE samples were created containing between 20 and 40 µg of protein (to compensate for variable cellular expression levels for various proteins of interest), 4X Laemmli Buffer (Bio-Rad #161-0747) containing 5% (v/v) 2-mercaptoethanol (Sigma-Aldrich M6250-250mL) and finally RIPA lysis buffer as filler to maintain a final volume of 50 µL for all samples. Samples were then loaded onto 10% Stain-Free acrylamide gels (Bio-Rad #161-0183) and run at 80 V for approximately 2 hours in tris-glycine SDS running buffer. The gels were then removed from their glass casing and activated for 1 minute under UV light (Bio-Rad Gel doc XR+). The samples were then transferred to low-flourescence PVDF membranes in tris-glycine transfer buffer (20% ethanol, 20% 5X tris-glycine, 60% water) at 30 V for 12 hours. Membranes were fluorescently scanned for total stain-free protein and incubated in blocking solution (PBS-T with 5% milk (w/v)) for 1 hour at room temperature followed by overnight incubation with primary antibody at the recommended dilution at 4°C. Membranes were then washed with PBS-T (1x PBS 0.05% (v/v) Tween-20) three times for 10 minutes each, followed by incubation with an IgG HRP-linked secondary antibody (Cell Signaling 7076S and 7074S) that was anti-mouse or anti-rabbit at 1:3000 dilution for 1 hour at room temperature. Membranes were again washed three times for 10 minutes each with PBS-T and finally ECL clarity (Bio-Rad #170-5060) was added after a 5 minute incubation autoradiography was performed using X-ray film (GE Healthcare) in a dark room with approximately 10 to 50 second exposures, depending on the expression level of the protein of interest.

Gene name	Disease name	Protein name	Antibody manufacturer	Catalogue number	Dilution
ADA	severe combined immunodeficiency disease (SCID)	Adenosine deaminase	Abcam	ab183631	1:3000
ALDH3A2	Sjogren-Larsson syndrome	Aldehyde dehydrogenase 3 family, member A2	Abcam	ab184171	1:2000
ASS1	citrullinemia	argininosuccinate synthase 1	Abcam	ab170900	1:1000
CAPN3	limb-girdle muscular dystrophies type 2A	Calpain 3	Abcam	ab51859	1:1000
CHM	choroideremia	Rab escort protein 1	Santa Cruz	sc-23905	1:500
DHCR7	Smith-Lemli-Opitz syndrome (SLOS)	7-dehydrocholesterol reductase	Abnova	H0000171 7-D01P	1:1000
FKRP	muscular dystrophy-dystroglycanopathy	Fukutin-Related Protein	Santa Cruz	sc-374642	1:100
GAA	Pompe's disease	acid alpha-glucosidase,	Abcam	Ab102815	1:1000
SGSH	Sanfilippo syndrome type A	Heparan Sulfate Sulfatase	Abcam	Ab96030	1:1000
UBB	Positive Control	Ubiquitin	Cell Signaling	3936S	1:1000

Table 2. Summary of primary antibodies used for the study.

2.6 Quantification

Quantification was performed by scanning the autoradiographs via a computer scanner (Canon LiDE at 600 dots per inch) and signal intensities were determined by measuring the area under the curve using Image J software (NIH). Loading controls were determined by measuring the area under the curve of the 10 most prominent stain-free protein peaks using Image Lab software (Bio-Rad).

2.7 Bortezomib Activity Verification by Ubiquitin Blotting

Approximately one out of every five Western blots (and at least one per every cell line) was blotted for ubiquitin to verify the efficacy of bortezomib. Patient cells were plated, treated and harvested per the aforementioned protocol. The sole variation is that the primary antibody was raised against ubiquitin (Cell Signalling 3936S) rather than a protein of interest.

2.8 AlamarBlue® Cell Viability Assay

Cell viability was assessed using the alamarBlue cell viability assay (ThermoFisher) where 20 µL of alamarBlue and 180 µL of aforementioned cell growth media were added in 96 well plates containing fibroblasts at approximately 70% confluency. Cells were then incubated at 37°C for approximately 6 hours until the media containing alamarBlue was reduced by the fibroblasts from a deep purple to a lighter pink, at which point the plates were measured by a spectrophotometer (Biotech synergy HT).

2.9 Statistical Analysis

For statistical analysis of protein level, the mean putative protein immunoblot band density was calculated (as the area under the curve determined by ImageJ software), normalized and the mean value is represented by the bar graphs in the figures. The added error bars represent the standard deviation, with the sample number (n-value) described in the figure caption.

CHAPTER 3

3.0 Results

3.1 Bortezomib Verification by Ubiquitin Immunoblot

In order to verify the correct concentration, potency and efficacy of bortezomib, a western blot with an antibody raised against ubiquitin was performed approximately every fifth western blot (Figures 2-5). Bortezomib-mediated 26S proteasome inhibition is reflected by a disproportionally large population of ubiquitin-tagged proteins with a correspondingly diverse range of sizes. Therefore an immunoblot against the ubiquitin tag itself should reveal a noticeable increase in the population of proteins and thus be represented as a large smear. This was in fact observed demonstrating the potency of the purchased bortezomib as well as its correct dilution, treatment regimen and penetration into the fibroblasts.

It was determined that the presence of bortezomib successfully increases the population of proteins tagged for ubiquitination that also remain undegraded. Interestingly, there does exist a small but noticeable population of tagged proteins of diverse sizes which are not degraded in the control cell lines. Furthermore despite the absence of a reliable means of quantification, the population of tagged proteins appears rather comparable between the 10 nM and 50 nM concentrations.

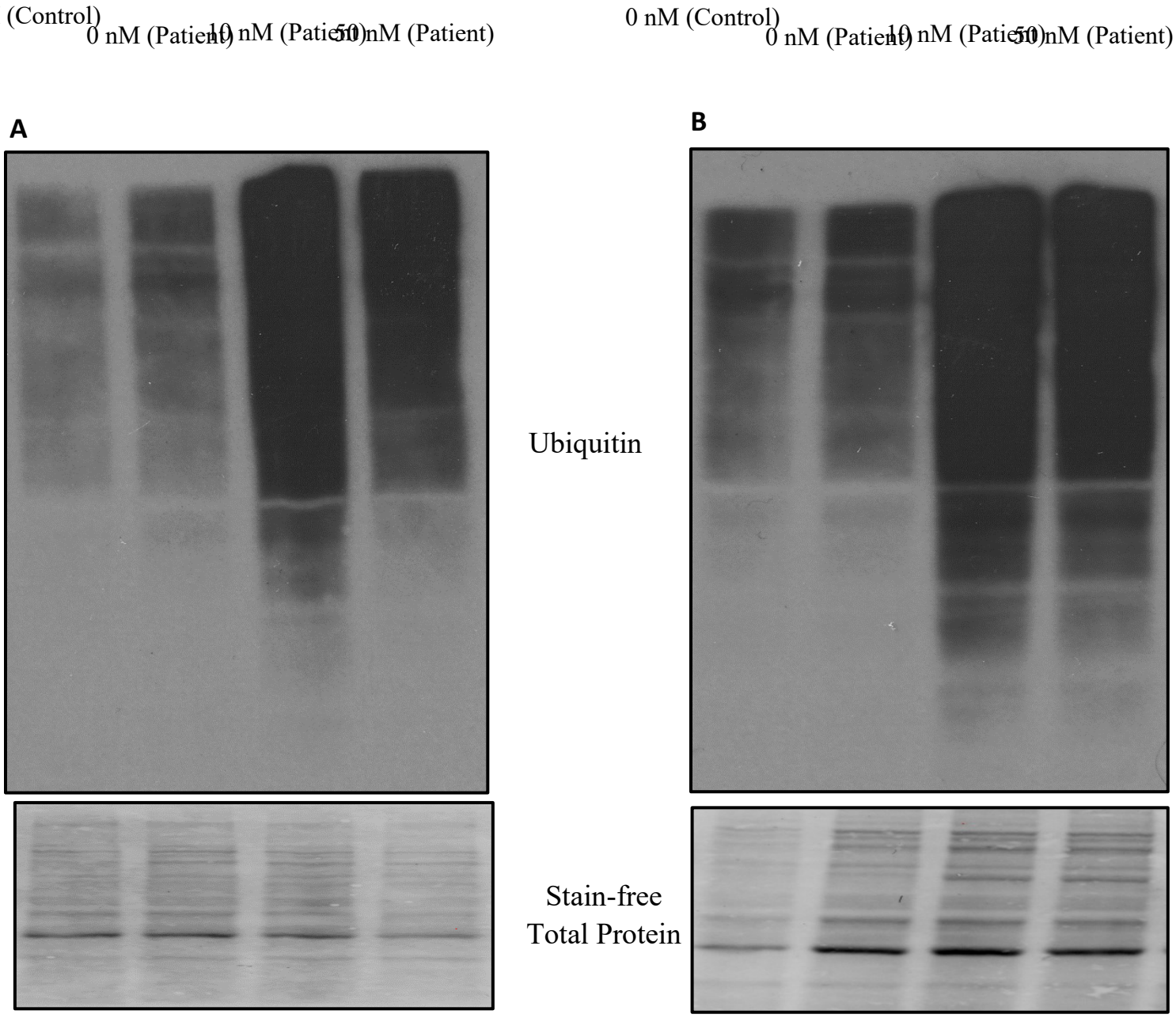


Figure 5. Bortezomib greatly increases the proportion of proteins that are tagged for ubiquitination but that are not ubiquitinated.
A) Bortezomib treatment of Muscular Dystrophy Limb-Girdle Type 2A (gene *CAPN3*) fibroblasts, Coriell patient GM03932, results in a significant global increase in the proportion of proteins tagged by ubiquitin that remain undegraded. Cells were treated for 24 hours.
B) Bortezomib treatment of Mucopolysaccharidosis Type IIIA (gene *SGSH*) fibroblasts, patient of (Muschol et al. 2011), results in a significant global increase in the proportion of proteins tagged by ubiquitin that remain undegraded. Cells were treated for 24 hours.

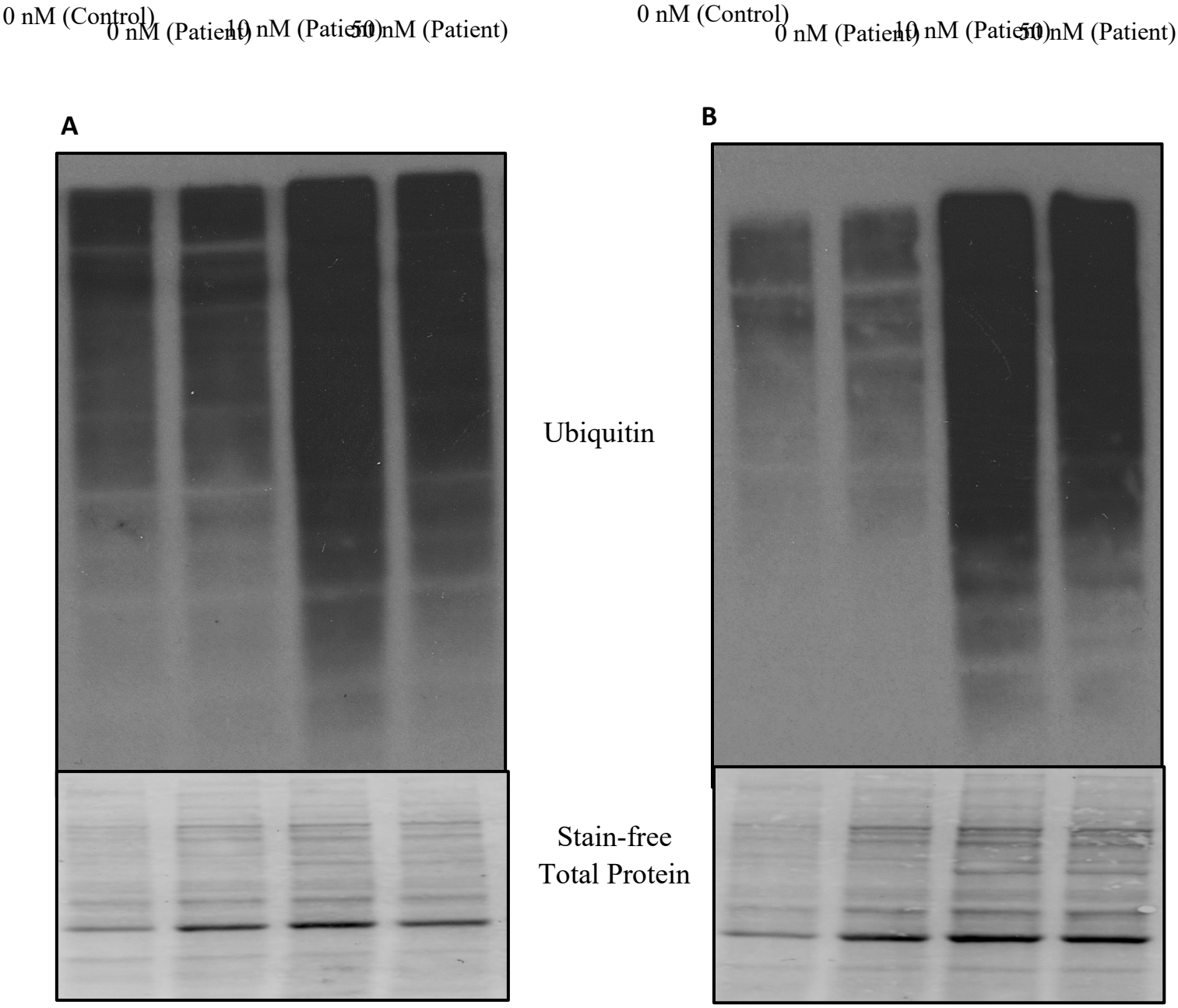


Figure 6. Bortezomib greatly increases the proportion of proteins that are tagged for ubiquitination but that are not ubiquitinated.
A) Bortezomib treatment of Severe combined immunodeficiency disease (gene *ADA*) fibroblasts, Coriell Cell patient GM02605, results in a significant global increase in the proportion of proteins tagged by ubiquitin to remain undegraded. Cells were treated for 24 hours.
B) Bortezomib treatment of Sjögren-Larsson syndrome (gene *ALDH3A2*) fibroblasts, Coriell Cell patient GM10641, results in a significant global increase in the proportion of proteins tagged by ubiquitin that remain undegraded. Cells were treated for 24 hours.

3.2 AlamarBlue® Cell Viability Assay

The next test to show the concentrations of bortezomib's suitability for the 24 hour treatment regimen used in this study and also provide some indication of bortezomib's potential as a long-term therapeutic is cell viability. This was performed by alamarBlue, a test based upon the non-fluorescent blue compound resazurin being reduced by natural cell activity to resorufin, a highly fluorescent compound, which readily absorbs 560 nm light and emits 590 nm light. The exact mechanism of action is not well understood, however the redox potential of the resazurin to resorufin conversion is 380 mV, well above that of the reactions in the electron transport chain. It is believed that cellular ETC activity can therefore supply sufficient energy for the reduction to resorufin; this reduction is therefore taken as a proxy for cell viability. This rate of conversion due to its fluorescent-emitting nature has a significant linear range and therefore the results obtained are proportional to the number of cells, and their overall health (Bio-Rad 2016).

It appears from (Figure 7, 8, 10) that 10 nM doses of bortezomib are fairly well tolerated, and that some patient cell lines tolerate higher doses for shorter durations better than a longer treatment regimen at a lower dose (Figure 9), and some patient cells the opposite (Figure 8). It is also surprising that healthy normal human fibroblasts are significantly more robust to bortezomib treatment (Figure 7) than patient cells, both in terms of higher doses and longer time periods, even though the patient cells, in theory, only have a single gene that is malfunctioning.

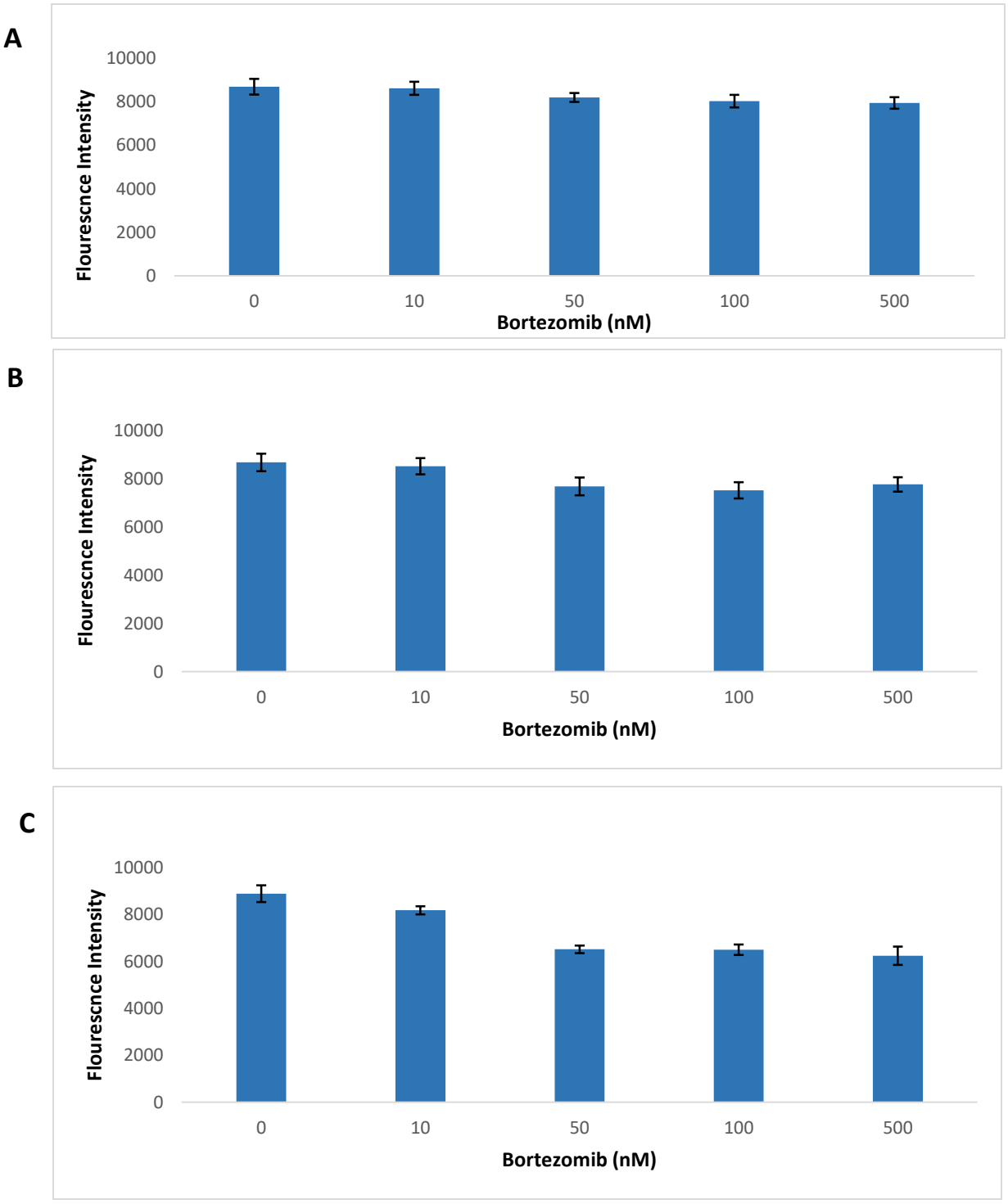
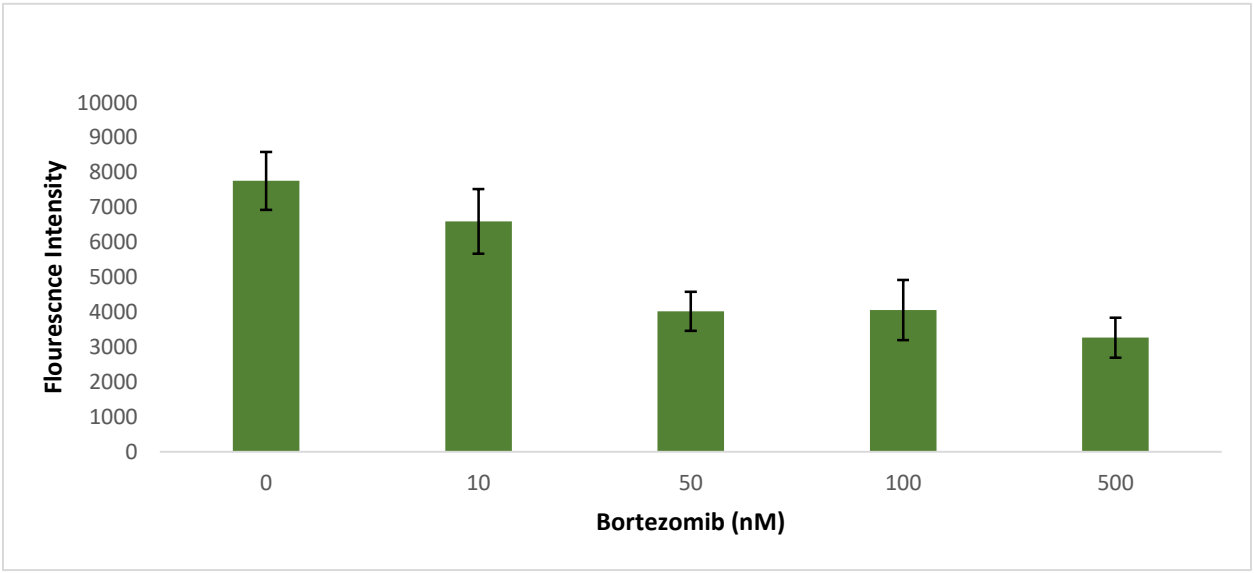


Figure 7. AlamarBlue shows bortezomib cell viability at 10 nM with a decrease in cell viability at 50 nM and even higher doses. Cell viability also decreases at 48 hr and 72 hr time points.

A) Normal human fibroblasts were treated with various doses of bortezomib for 24 hours
B) Normal human fibroblasts were treated with various doses of bortezomib for 48 hours
C) Normal human fibroblasts were treated with various doses of bortezomib for 72 hours
(results are presented as the mean ± standard deviation, for n = 6)

A



B

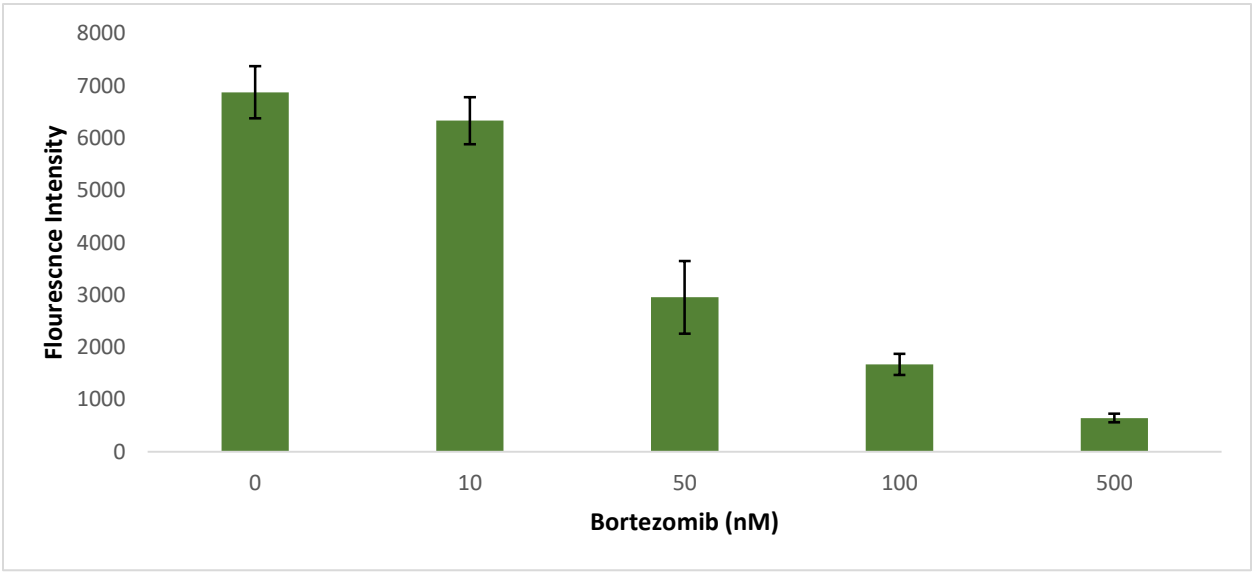


Figure 8. AlamarBlue shows bortezomib cell viability at 10 nM with a significant decrease in cell viability at 50 nM and even higher doses. Cell viability also sharply declines at 48 hr.
A) FKRP patient fibroblasts (GM23868) were treated with various doses of bortezomib for 24 hours
B) FKRP patient fibroblasts (GM23868) were treated with various doses of bortezomib for 48 hours
(results are presented as the mean ± standard deviation, for n = 6)

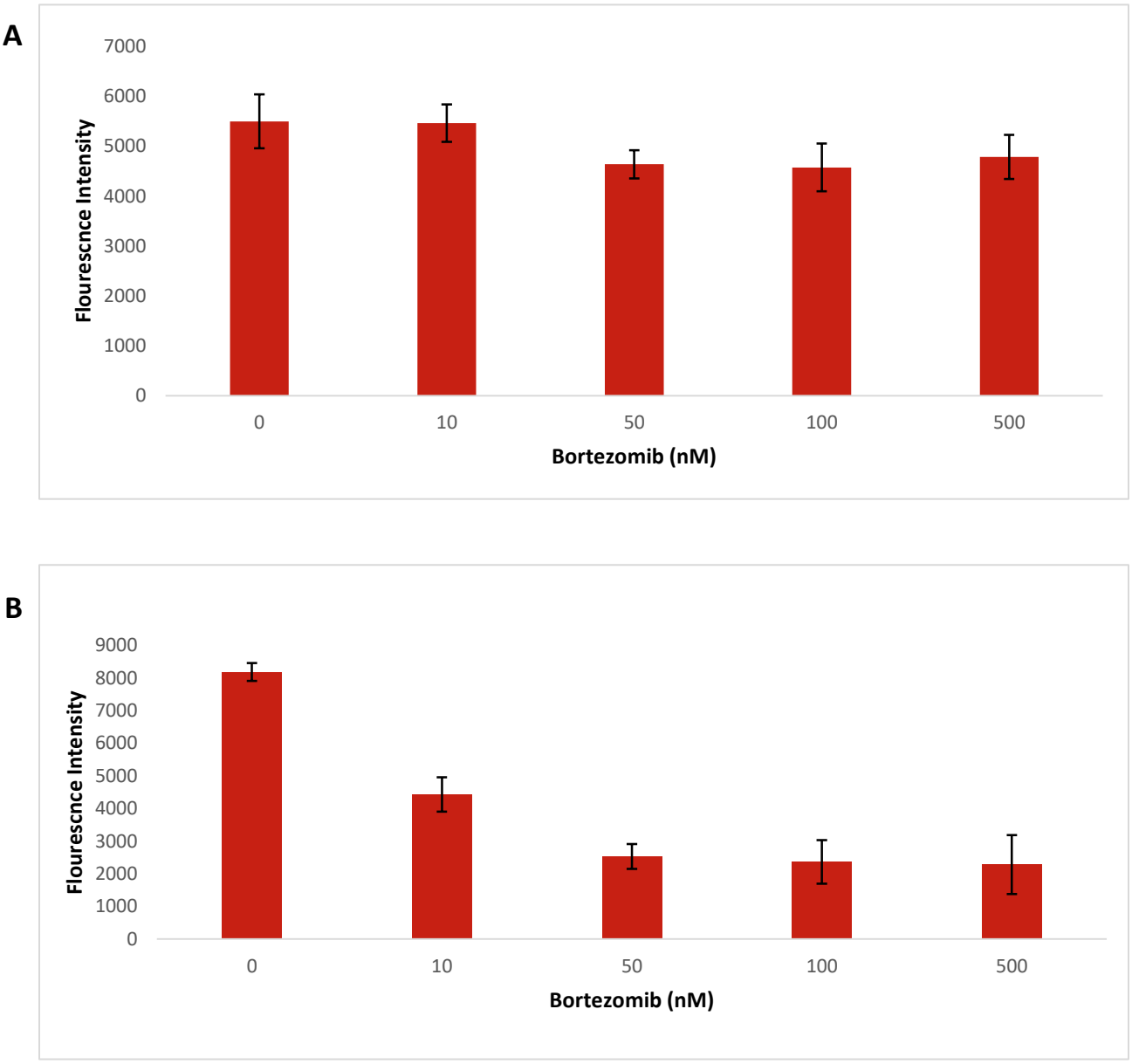


Figure 9. AlamarBlue shows bortezomib cell viability at 10 nM with only a slight decrease in cell viability at 50 nM and higher doses, at 24 hr. Cell viability sharply declines at 48 hrs for all doses.

A) PMM2 patient fibroblasts (GM20948) were treated with various doses of bortezomib for 24 hours

B) PMM2 patient fibroblasts (GM20948) were treated with various doses of bortezomib for 48 hours

(results are presented as the mean ± standard deviation, for n = 6)

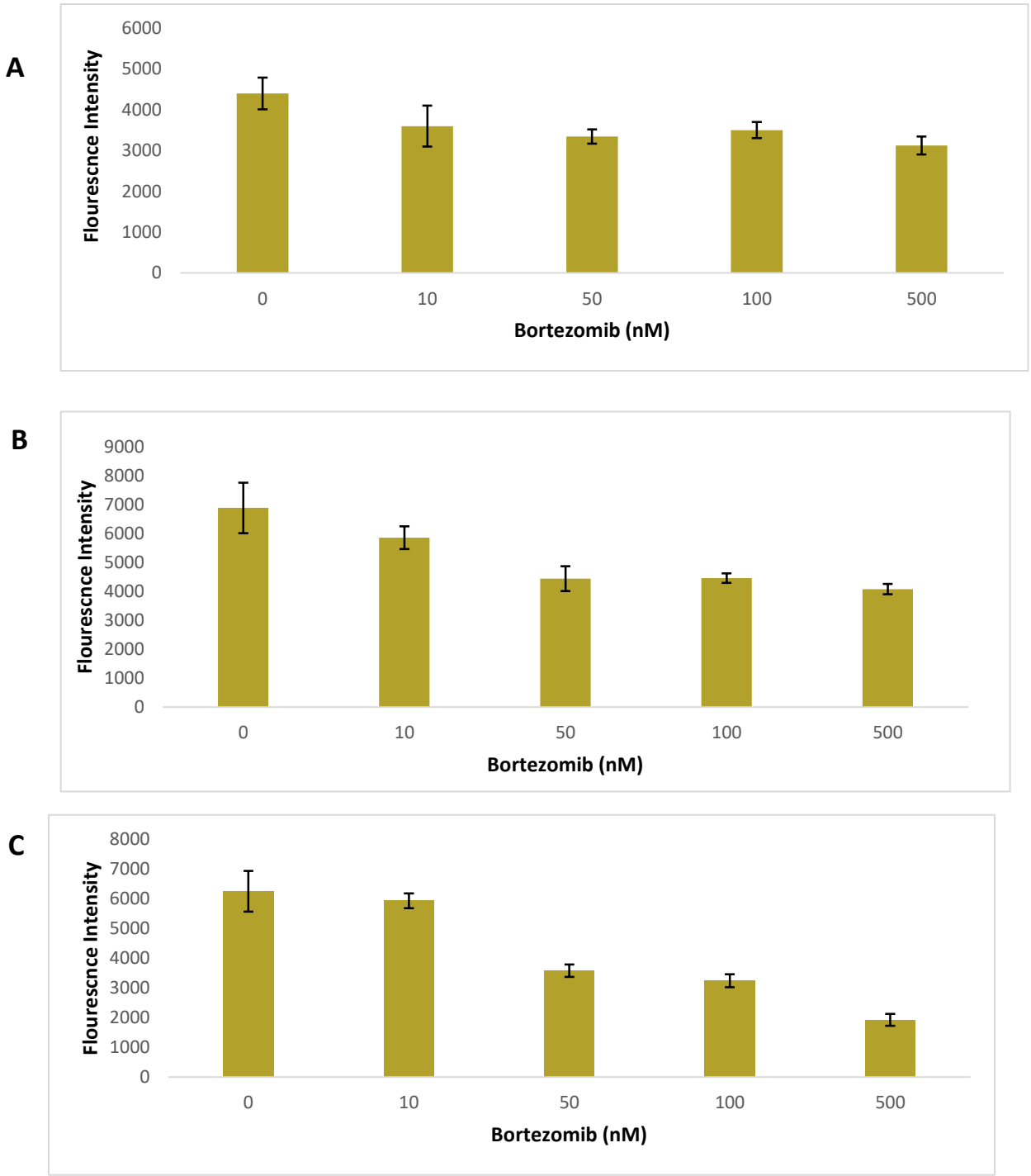


Figure 10. AlamarBlue shows poor bortezomib cell viability at any dose, with additional time points reducing cell viability even further. The cells are from a local D-bifunctional protein deficiency (DBP) patient.

A) patient fibroblasts (DBP) were treated with various doses of bortezomib for 24 hours
B) patient fibroblasts (DBP) were treated with various doses of bortezomib for 48 hours
C) patient fibroblasts (DBP) were treated with various doses of bortezomib for 72 hours
(results are presented as the mean ± standard deviation, for n = 6)

3.3 Western Blotting by Stain-Free

3.3.1 Treatments with 10 and 50 nM at 24 hours

The impact of bortezomib on mutated protein level was assessed by Western blot with 24 hours chosen as the primary treatment duration as this was the duration of all three cases of bortezomib's success in the literature (Shimada et al. 2011; Azakir et al. 2012; Macías-Vidal et al. 2014). The concentrations of 10 and 50 nM were chosen for similar reasons, with 10 nM being a low but effective dose (Shimada et al. 2011; Azakir et al. 2012), and 50 nM being on the upper range of cell viability and beyond the required dose for visible efficacy for any treatment found in the literature (Shimada et al. 2011; Azakir et al. 2012; Macías-Vidal et al. 2014).

Bortezomib did not upregulate any of the mutated protein in any of the patient cell lines attempted. Compromises were made with respect to working entirely on missense mutated proteins, as the selection and availability of patient cells that were genotyped and confirmed to be missense was too scarce for the scope of this study. Cell types were used with an unknown genotype, multiple amino acid substitutions, or even a nonsense mutation but with the vast majority of the protein in-tact. These were selected as there was a real dearth of missense cell lines.

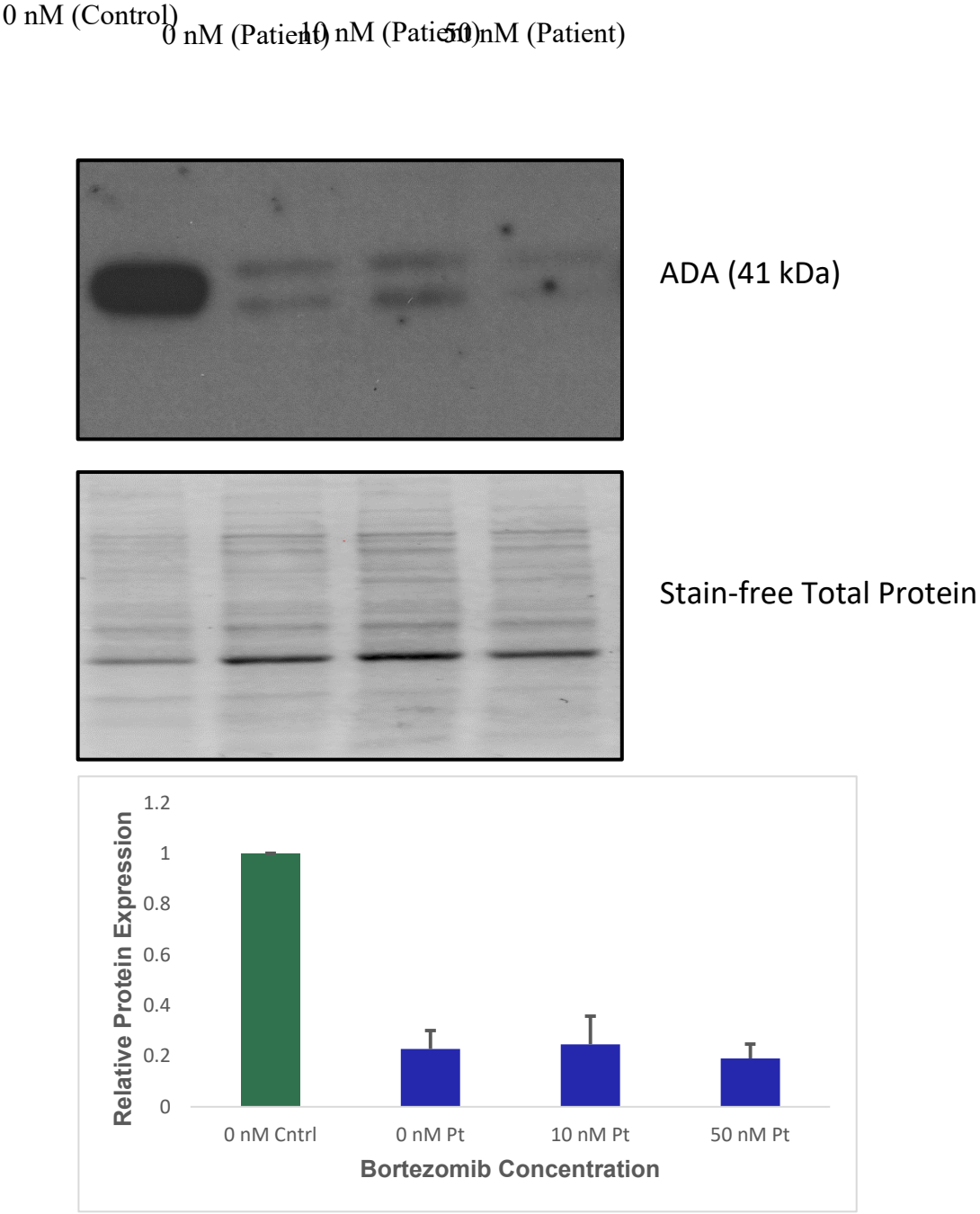


Figure 11. Patient ADA levels are much lower than wild type and are unaffected by bortezomib treatment.
Representative western blot quantifying ADA levels in fibroblasts for a control and a severe combined immunodeficiency disease (SCID) patient treated with bortezomib for 24 hrs, in compound heterozygote SCID fibroblasts with the following mutation: Arg101Trp/Arg211His; Coriell GM02605, 30 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

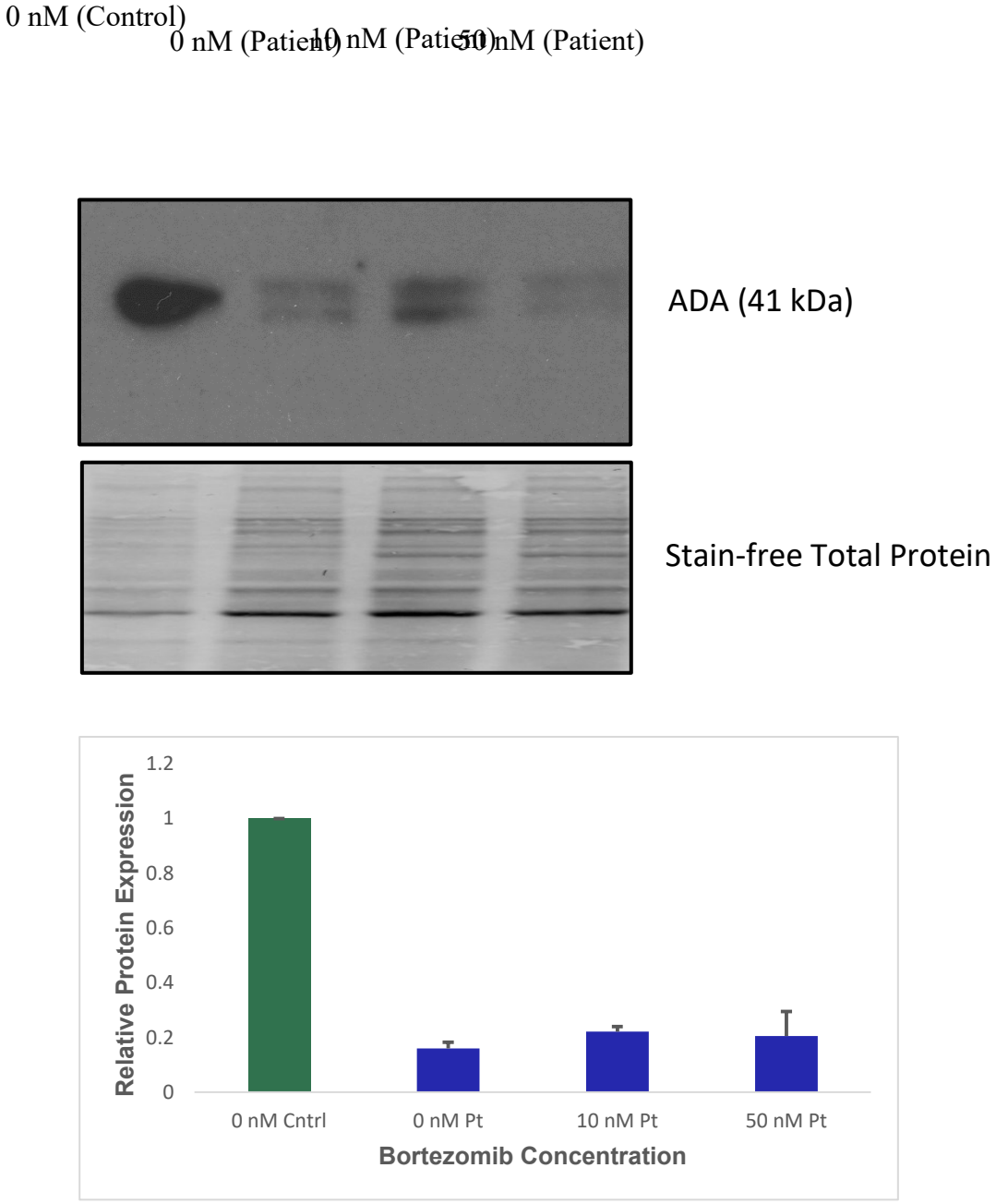


Figure 12. Patient ADA levels are much lower than wild type and are unaffected by bortezomib treatment.
Representative western blot quantifying ADA levels in fibroblasts for a control and a severe combined immunodeficiency disease (SCID) patient treated with bortezomib for 24 hrs, in compound heterozygote SCID fibroblasts with the following mutation: Leu107Pro/Arg211Cys; Coriell GM04395, 30 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

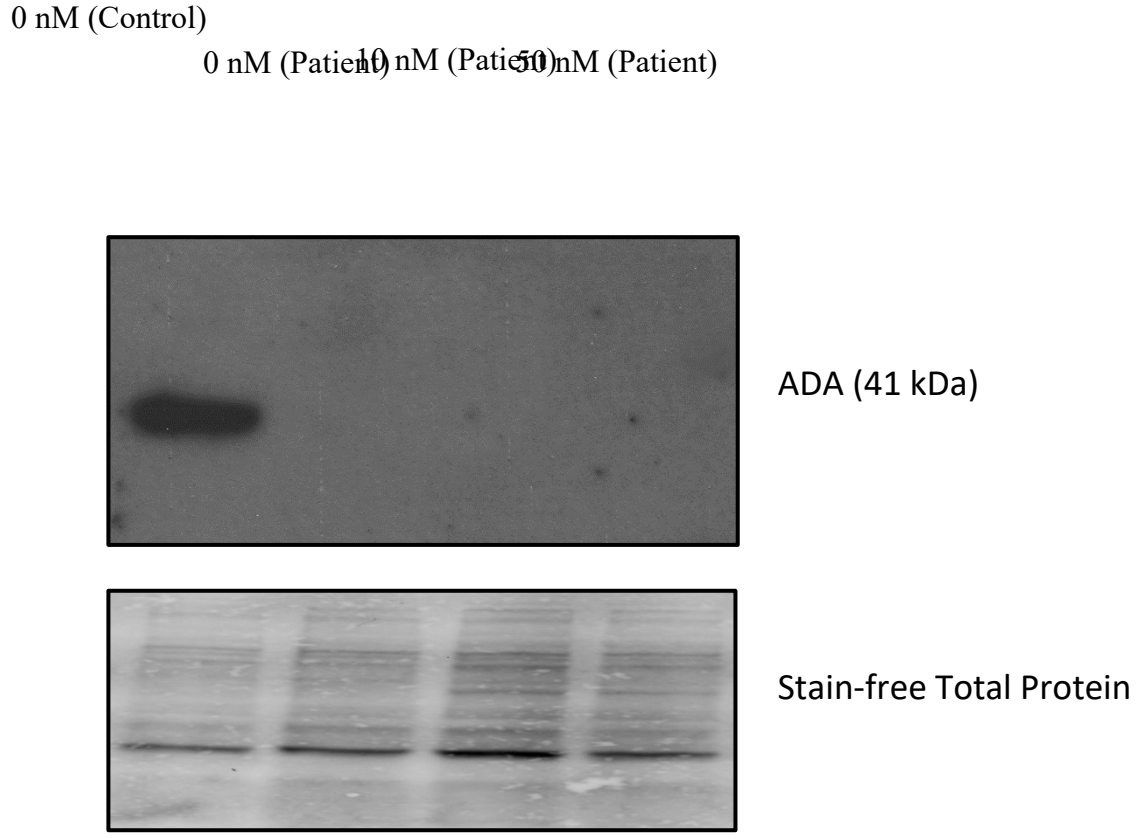


Figure 13. Patient ADA levels are too low to be detectible and appear unaffected by bortezomib treatment.
Representative western blot quantifying ADA levels in fibroblasts for a control and a severe combined immunodeficiency disease (SCID) patient treated with bortezomib for 24 hrs, in compound heterozygote SCID fibroblasts with the following mutation: A329V/exon 4 deletion; Coriell GM02824, 30 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

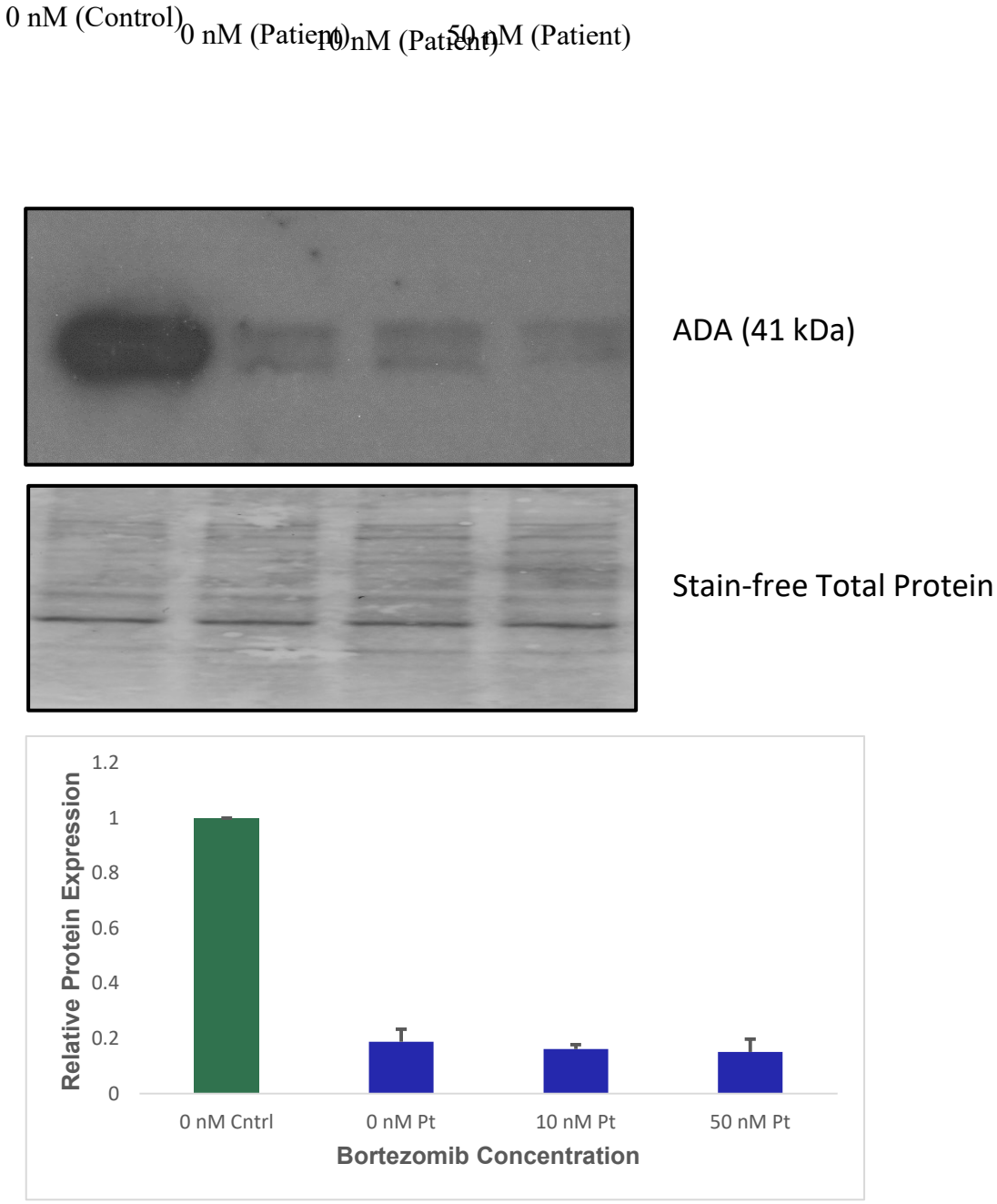


Figure 14. Patient ADA levels are much lower than wild type and are unaffected by bortezomib treatment.
Representative western blot quantifying ADA levels in fibroblasts for a control and a severe combined immunodeficiency disease (SCID) patient treated with bortezomib for 24 hrs, in compound heterozygote SCID fibroblasts with the following mutation: Arg101Gln/IVS1(+1GT>CT (resulting in unstable mRNA); Coriell GM02445, 30 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

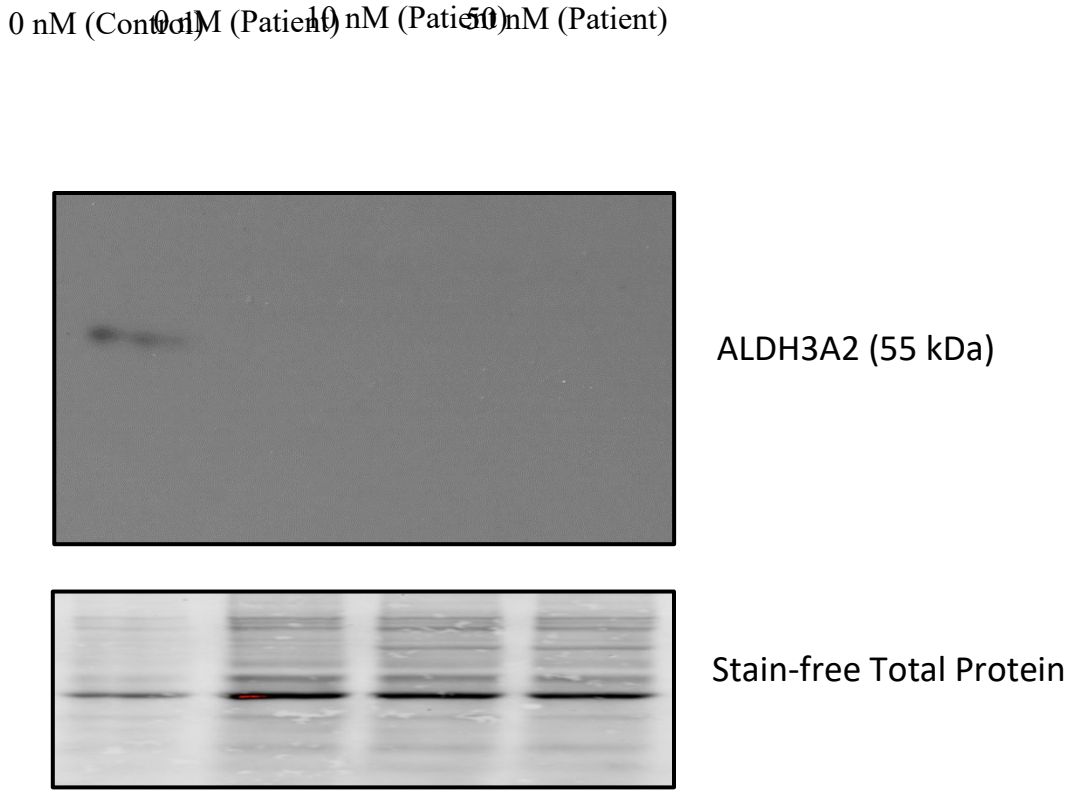


Figure 15. Patient ALDH3A2 levels are too low to be detectable and appear unaffected by bortezomib treatment.
Representative western blot quantifying ALDH3A2 levels in fibroblasts for a control and a Sjögren-Larsson syndrome patient treated with bortezomib for 24 hrs, in premature termination at amino acids 434 and 457 fibroblasts with 7% biochemical activity; Coriell GM10640, 30 μg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

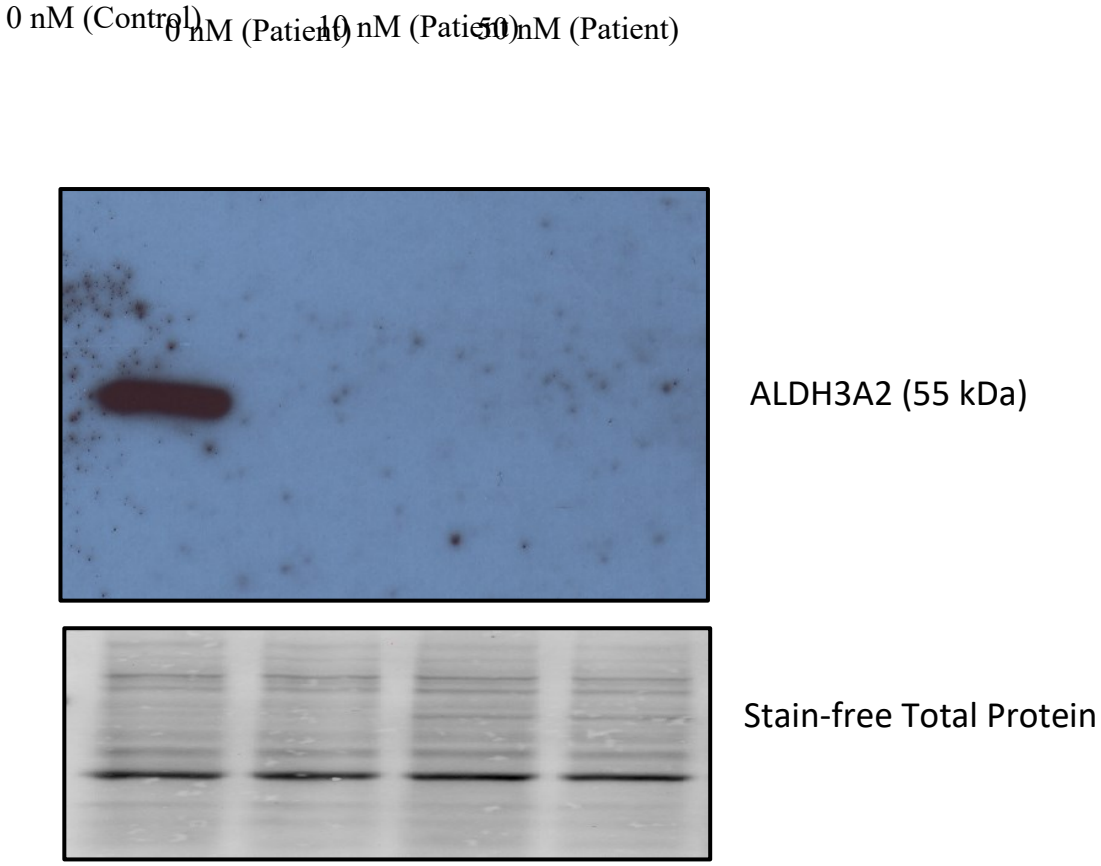


Figure 16. Patient ALDH3A2 levels are too low to be detectable and unaffected by bortezomib treatment.
Representative western blot quantifying ALDH3A2 levels in fibroblasts for a control and a Sjögren-Larsson syndrome patient treated with bortezomib for 24 hrs, in cells with the mutation of Ala314Gly/Pro315Ala with six amino acids added between the pair Ala-Lys-Ser-Thr-Val-Gly and with 5% biochemical activity, Coriell GM10641, 30 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

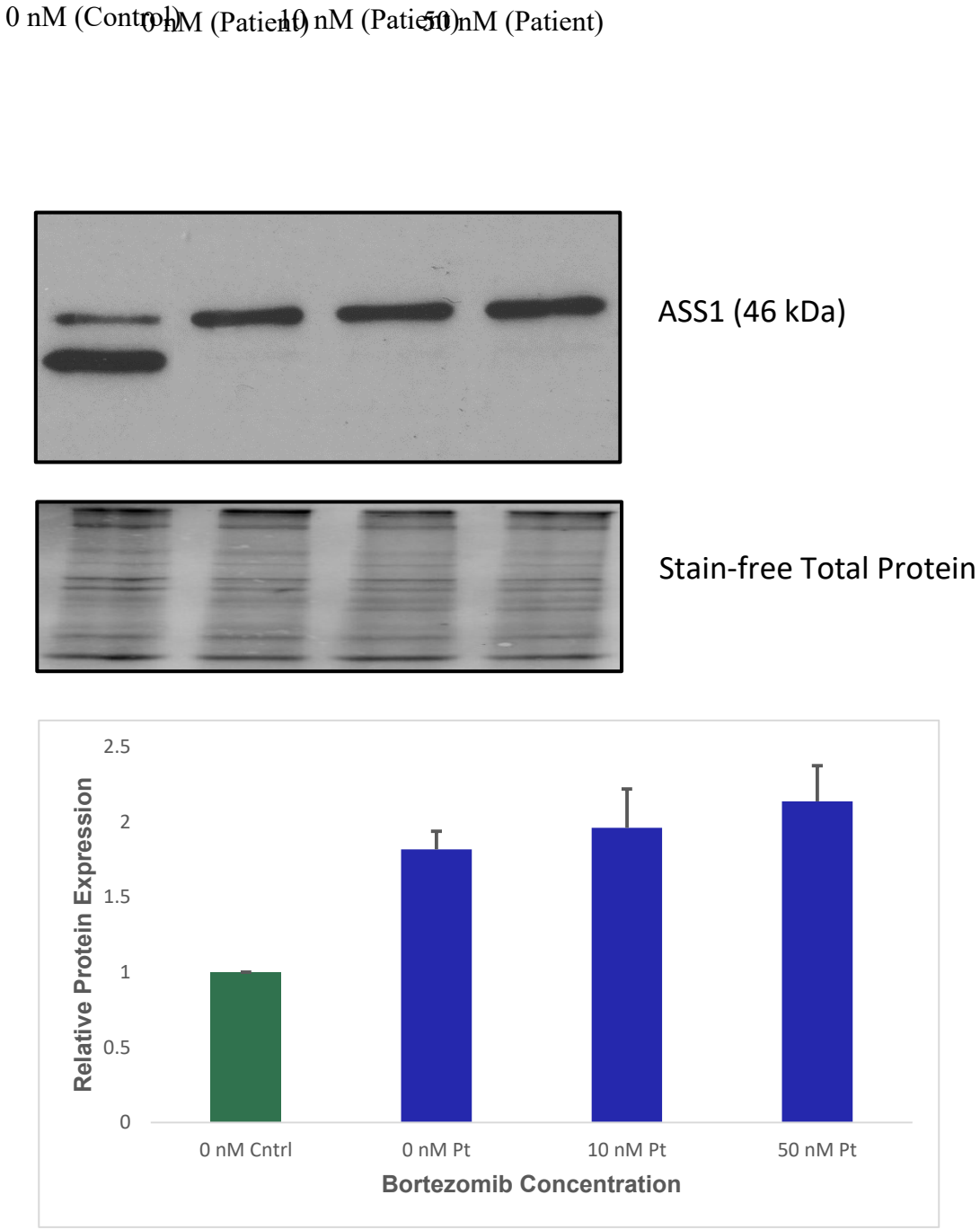


Figure 17. Patient levels of ASS1 are significantly higher than wild type levels, but are unaffected by bortezomib.
Representative western blot quantifying ASS1levels in fibroblasts for a control and a Citrullinemia patient treated with bortezomib for 24 hrs, in cells with an unknown mutation and 38% biochemical activity; Coriell GM01058, 40 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

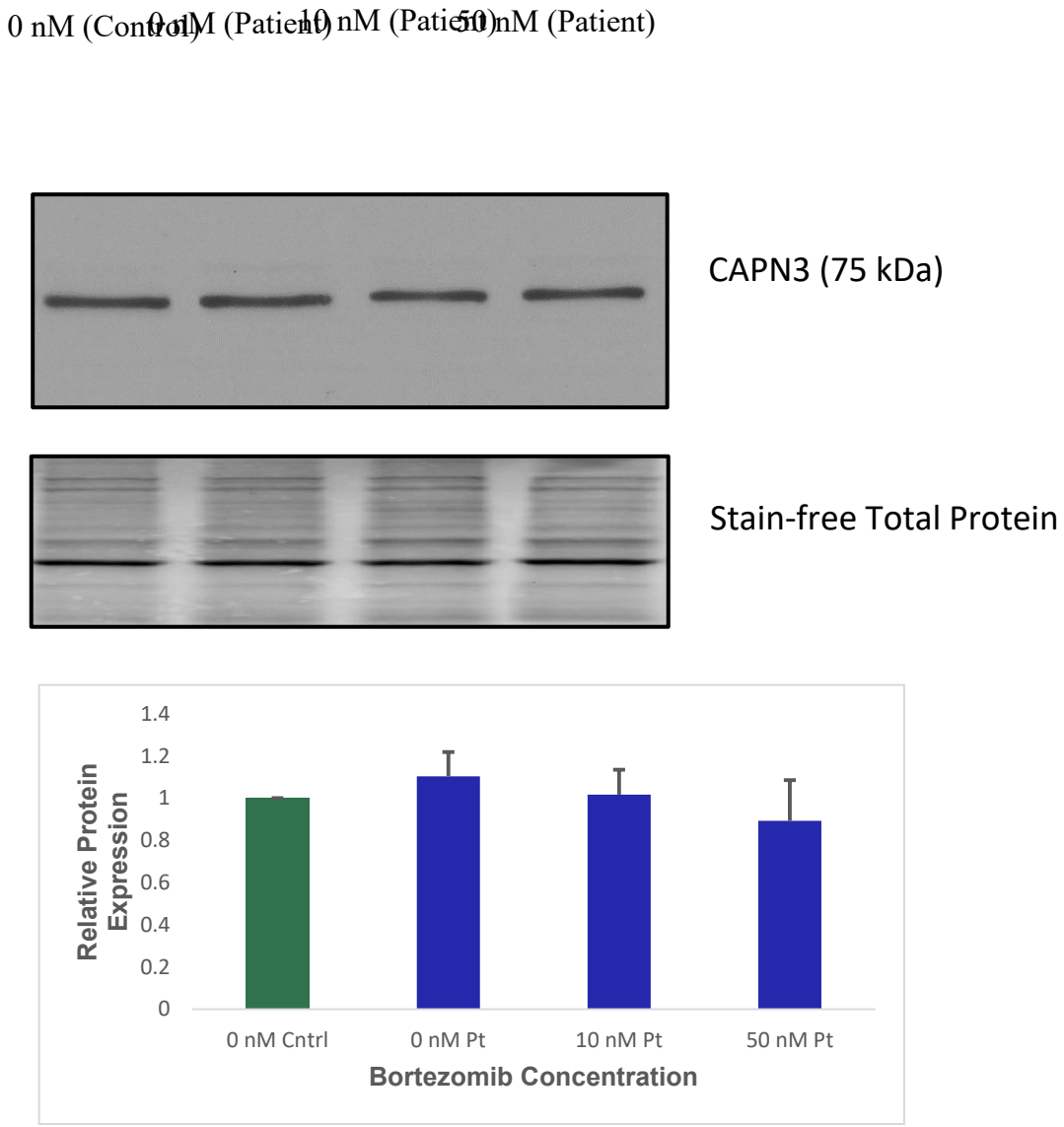


Figure 18. Patient levels of CAPN3 are very similar to wild type levels, and are unaffected by bortezomib.
Representative western blot quantifying CAPN3 levels in fibroblasts for a control and a Limb-Girdle Type 2A Muscular Dystrophy patient treated with bortezomib for 24 hrs, in cells with an unknown mutation; Coriell GM03932, 15 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

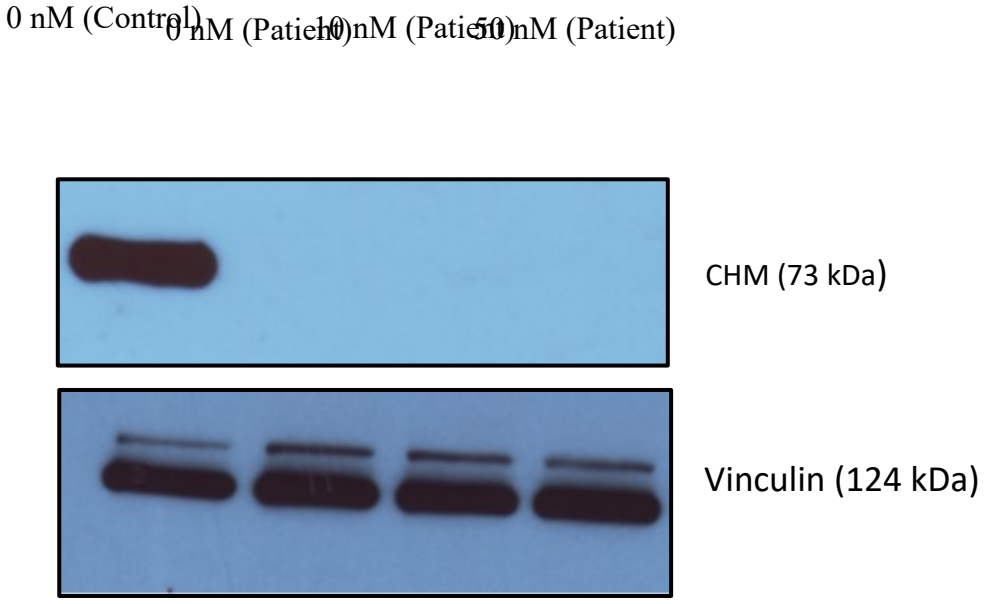


Figure 19. Patient CHM levels are too low to be detectable and unaffected by bortezomib treatment.
Representative western blot quantifying CHM levels in fibroblasts for a control and a choroideremia patient treated with bortezomib for 24 hrs, in cells with an unknown mutation; Coriell GM06940, 30 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

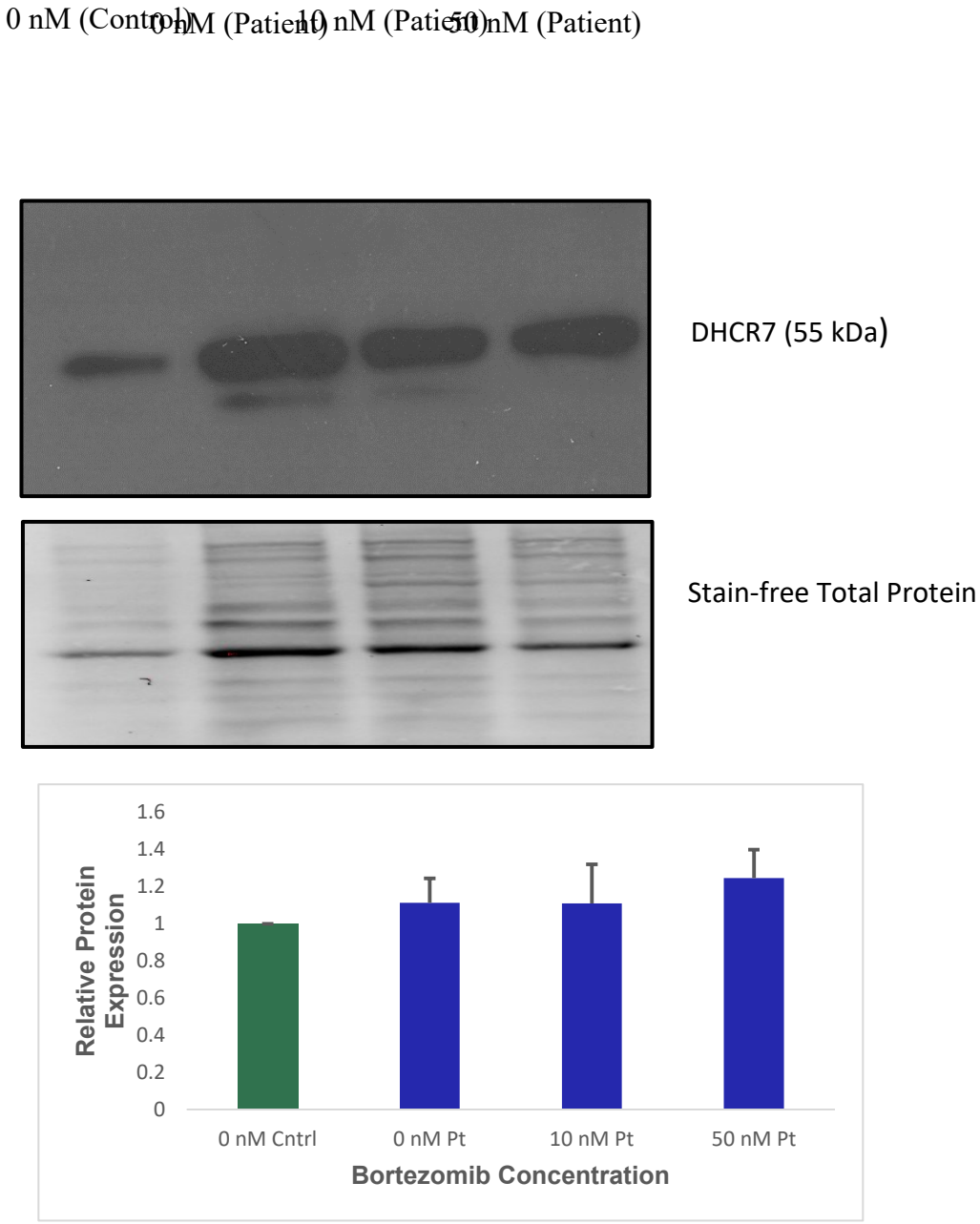


Figure 20. Patient levels of DHCR7 are very similar to wild type levels, and are unaffected by bortezomib.
Representative western blot quantifying DHCR7 levels in fibroblasts for a control and a Smith-Lemli-Opitz syndrome patient treated with bortezomib for 24 hrs, in cells with a translocation breakpoint at chromosome 13 and chromosome 14; Coriell GM03044, 30 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

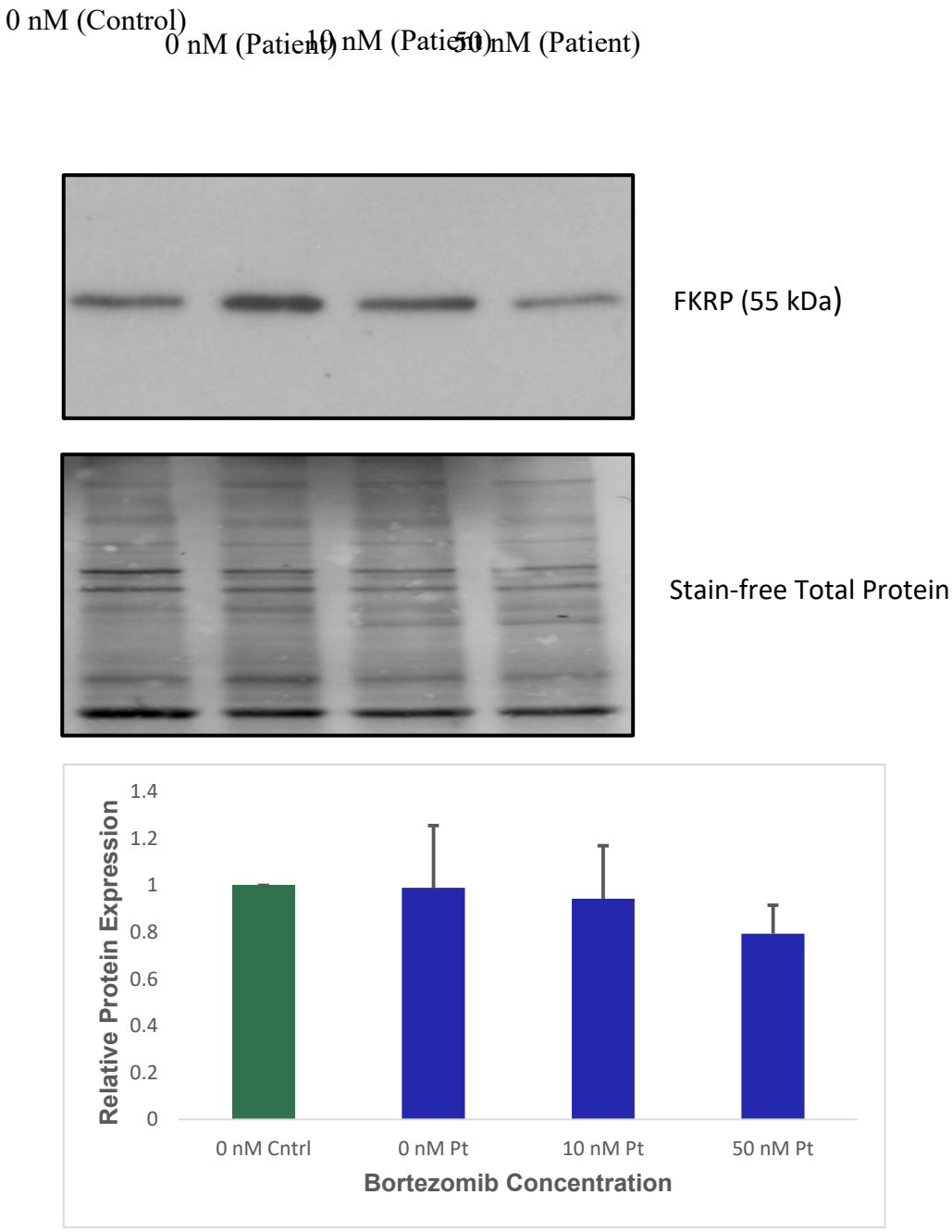


Figure 21. Patient levels of FKRP are very similar to wild type levels, and are unaffected by bortezomib.
Representative western blot quantifying FKRP levels in fibroblasts for a control and a Limb-Girdle Type 5C Muscular Dystrophy patient treated with bortezomib for 24 hrs, in cells with a homozygous mutation for Leu276Ile and unknown biochemical activity; Coriell GM23868, 50 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

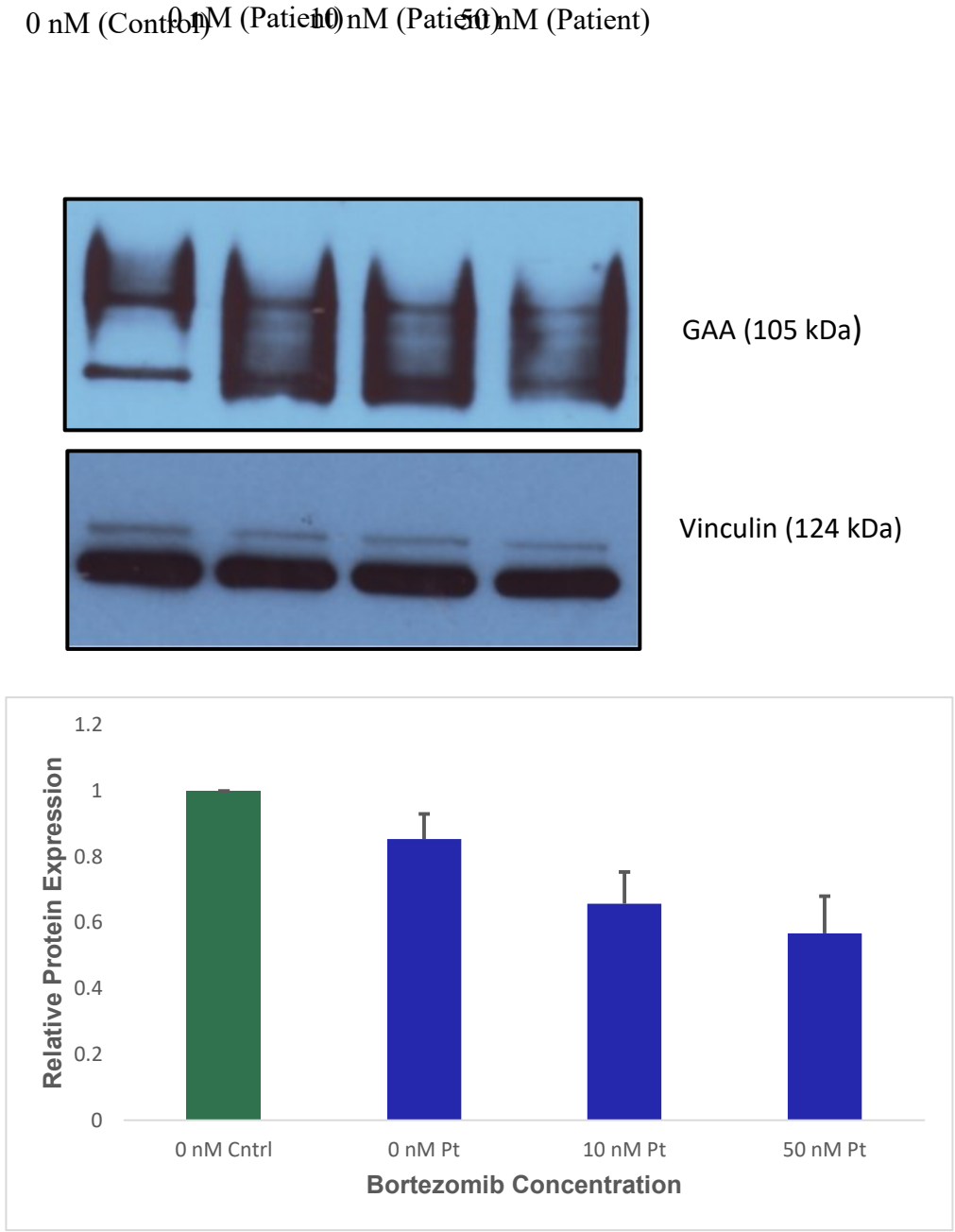


Figure 22. Patient levels of GAA do not appear to be treated by bortezomib, however accurate quantification is impossible due to dumbbell shaped bands.
Representative western blot quantifying GAA levels in fibroblasts for a control and a glycogen storage disease II patient treated with bortezomib for 24 hrs, in cells of a Canadian patient that is homozygous for the mutations c.-45T>G; and c510T>C, 30 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 3)

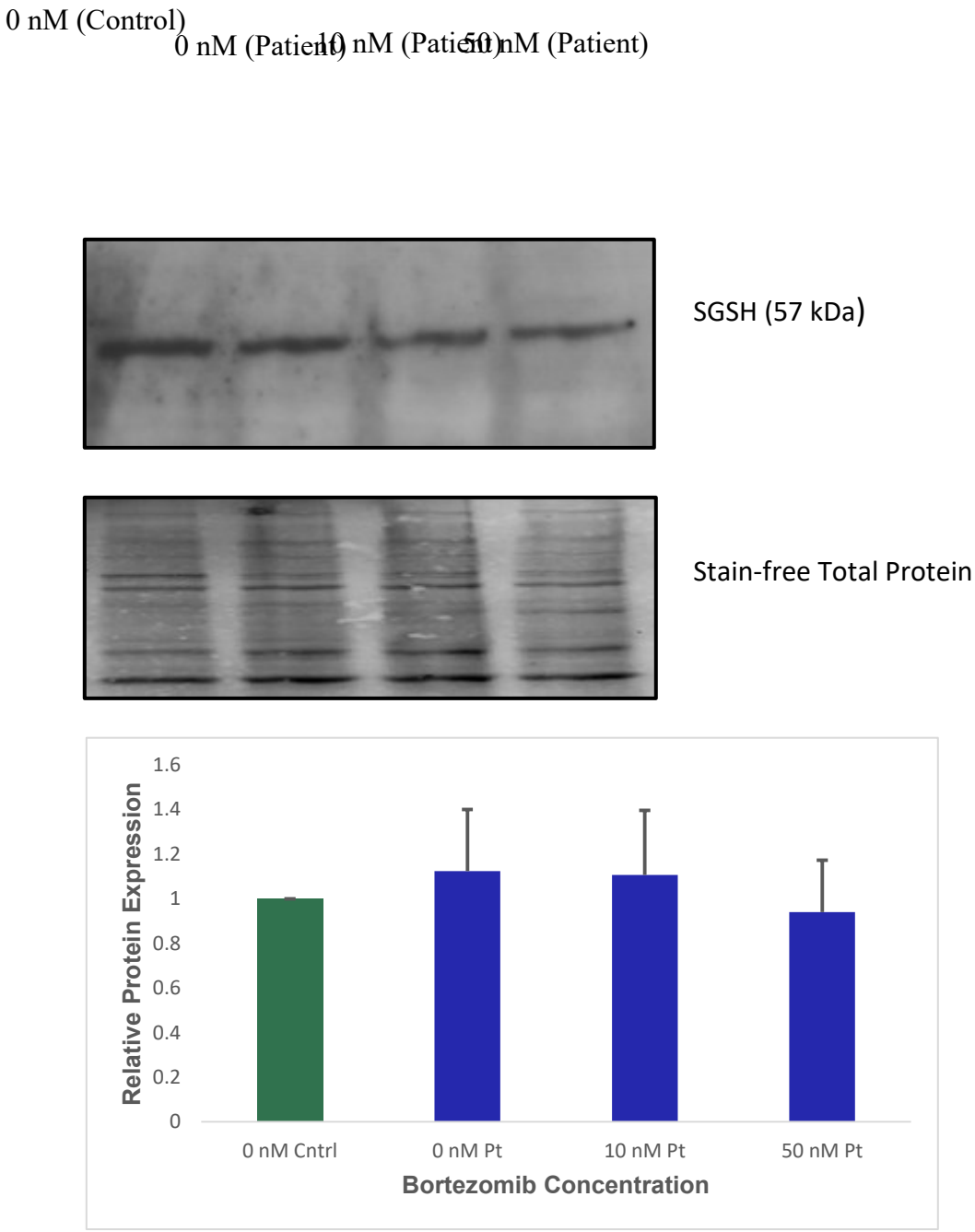


Figure 23. Patient levels of SGSH are very similar to wild type levels, and are unaffected by bortezomib.
Representative western blot quantifying SGSH levels in fibroblasts for a control and a mucopolysaccharidosis Type IIIA patient treated with bortezomib for 24 hrs, in cells of a compound heterozygote for S298P and R74C and of unknown biochemical activity, that is a patient of (Muschol et al. 2011), 30 µg of protein loaded in each lane. (results are presented as the mean ± standard deviation, for n = 3)

3.3.2 Treatments with 10 and 50 nM at 48 hours

The increase in treatment exposure to 48 hours was performed for rigour, as every case of successful bortezomib treatment in the literature was performed at 24 hours. There was unfortunately no difference in bortezomib protein upregulation at 48 hours compared to 24 hours for the small subset of patient cell lines treated, and the experiments were performed merely in duplicate, as opposed to in triplicate, for expediency.

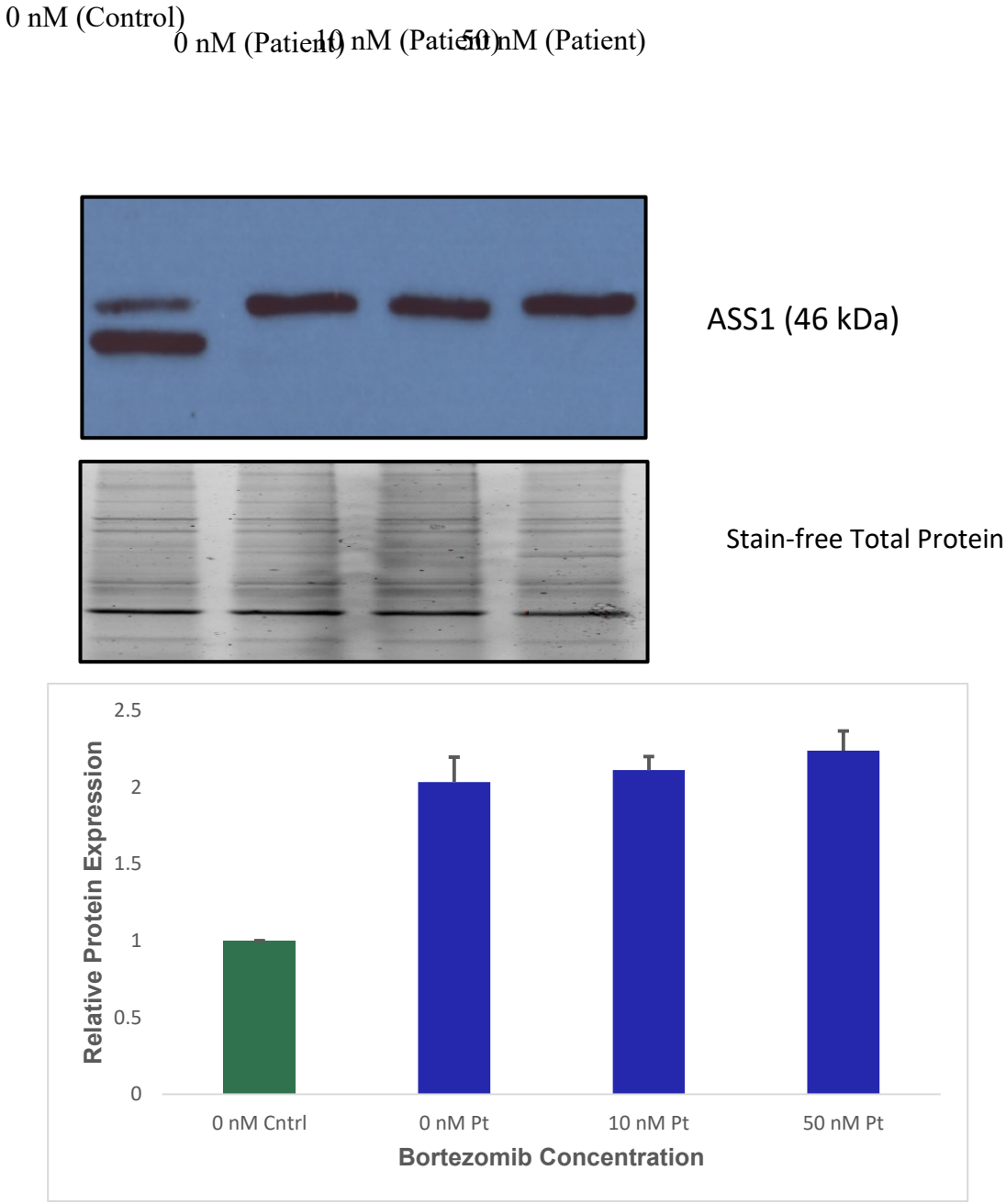


Figure 24. Patient levels of ASS1 are significantly higher than wild type levels, but are unaffected by bortezomib.
Representative western blot quantifying ASS1levels in fibroblasts for a control and a Citrullinemia patient treated with bortezomib for 48 hrs, in cells with an unknown mutation and 38% biochemical activity; Coriell GM01058, 40 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 2)

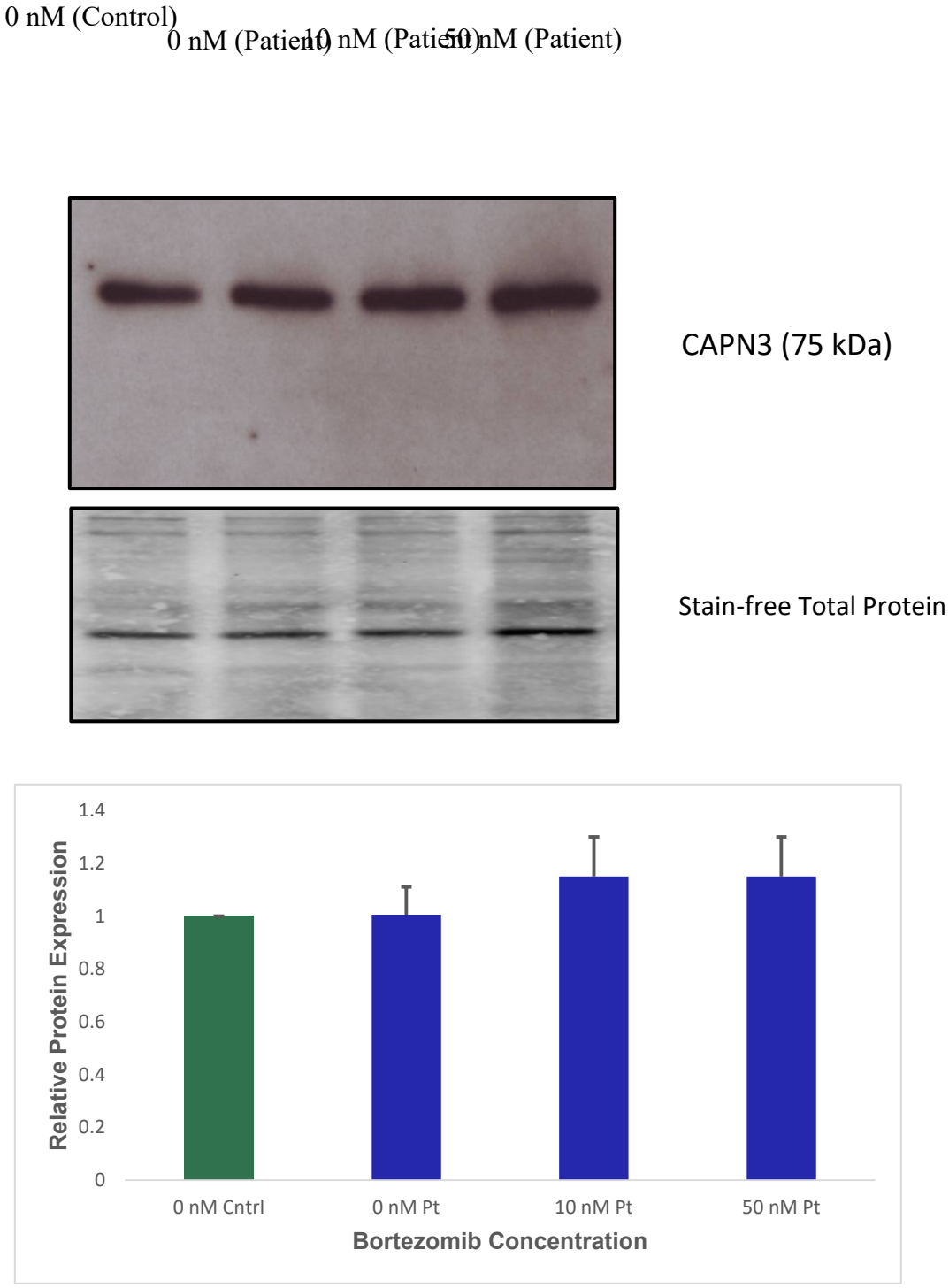


Figure 25. Patient levels of CAPN3 are very similar to wild type levels, and are unaffected by bortezomib.
Representative western blot quantifying CAPN3 levels in fibroblasts for a control and a Limb-Girdle Type 2A Muscular Dystrophy patient treated with bortezomib for 48 hrs, in cells with an unknown mutation; Coriell GM03932, 15 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 2)

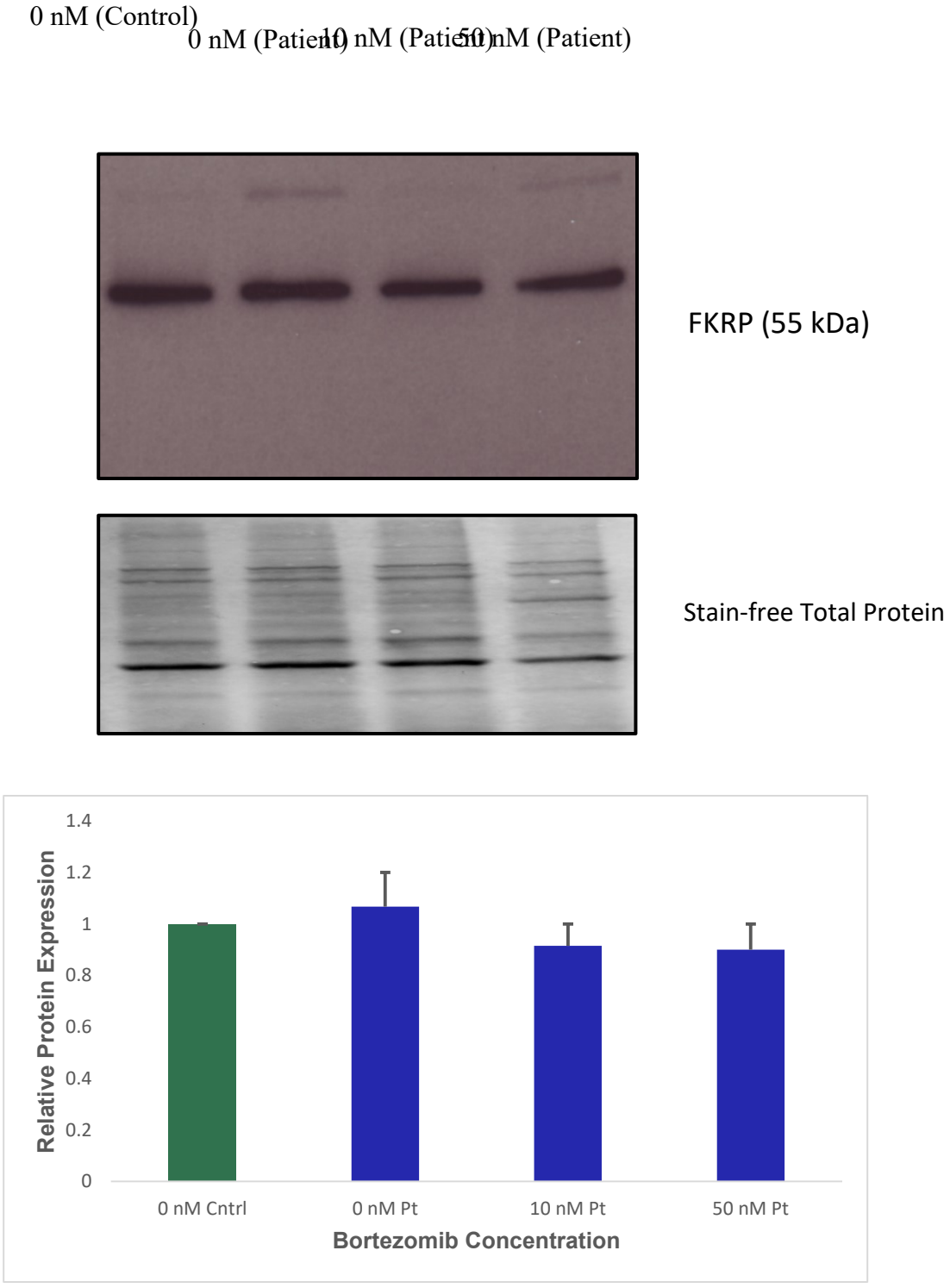


Figure 26. Patient levels of FKRP are very similar to wild type levels, and are unaffected by bortezomib.
Representative western blot quantifying FKRP levels in fibroblasts for a control and a Limb-Girdle Type 5C Muscular Dystrophy patient treated with bortezomib for 48 hrs, in cells with a homozygous mutation for Leu276Ile and unknown biochemical activity; Coriell GM23868, 50 µg of protein loaded in each lane (results are presented as the mean ± standard deviation, for n = 2)

CHAPTER 4

4.0 Discussion

4.1 Major Findings

Bortezomib treatment under the conditions of this study had no effect on mutated protein levels in 13 cell lines representing 9 disorders (Figures 11-23), despite the great diversity in recessive disease type, protein subcellular localization and patient protein level. Although the lack of bortezomib mediated protein upregulation was observed consistently across all cell lines, the underlying reasons may be themselves diverse and potentially explainable. Bortezomib (10 nM) penetrated patient cells and inhibited proteasome activity (Figure 5 and 6) and fibroblast viability was not adversely affected in most of the cell lines even after multiple days of exposure (Figures 7 to 10); therefore poor drug bioavailability, efficacy or significant cytotoxicity can be ruled out as underlying reasons.

4.1.1 Patient Cell Lines with No Baseline Expression

A number of cell lines which were selected in the absence of protein expression data from Coriell had mutated protein levels below the limit of detection (Figures 13, 15, 16, 19). Even should bortezomib had been effective in increasing the amount of protein several-fold, it is highly likely the protein would still remain so low in quantity as to be undetectable and thus unlikely to improve the disease phenotype. In keeping with this, there is no example in the literature of proteasomal inhibition raising an undetected protein above the level of detection. ALDH3A2 mutant cell lines (Figure 15) segregated nonsense mutations at amino acids 434 and 457, we would not anticipate proteasomal inhibition to increase the protein. However due to the full length protein being only 485 amino acids and the functional domain ending at amino acid 429 (EMBL-EBI 2016), in the absence of more promising cell lines, we speculated there might be a functional truncated protein. ALDH3A2 mutant cell lines (Figure 16) featured a more

complex mutation as they included a six amino acid addition in one of the alleles with unknown implications although 5% residual enzymatic activity was reported on the Coriell web page. Mutant CHM (Figure 19) featured an unknown mutation and biochemical activity, which is likely a premature stop mutation but was the only cell line available for this recessive disorder.

4.1.2 Patient Cell Lines with Low Baseline Expression

Disease genes (Figures 11, 12 & 14) showed low levels of protein prior to, and after, bortezomib treatment. Low protein levels are often observed in missense mutations (Cartegni et al. 2002). Due to normal levels of transcription and translation, once the protein is synthesized, its thermodynamic instability is detected and it's targeted for degradation by the proteasome. These candidate disorders would be anticipated to have the greatest likelihood of responding to bortezomib treatment. One possible explanation for the lack of bortezomib treatment efficacy might be transcriptional instability resulting in the reduction of protein. This is emphasized by Coriell in the case of a mutant ADA cell line (Figure 14) which contains one alleles that is mentioned as having unstable mRNA. However the other five alleles across the three mutant ADA cell lines tested are simple missense mutations which should be less likely to lead to mRNA instability (Belgrader & Maquat 1994). Another possibility might be a reduced level of translation, where the mRNA is stable but due to the missense mutation itself, the mRNA is transcribed at a much lower level, resulting in the low observed levels of protein and again not receptive to bortezomib treatment.

4.1.3 Patient Cell Lines with Full Baseline Expression

Figures 17, 18, 20, 21, 22 and 23 all showed essentially wild type levels of the diseased protein, this is possibly due to the missense mutated proteins not being targeted for degradation but having a poor level of function, perhaps due to the mutation being on the functional domain,

but not destabilizing the tertiary structure. The genotype of the mutations are unknown for (Figures 17, 18, 21) and the cell line shown in (Figure 20) features a chromosomal translocation and the cell line in (Figure 22) features a splicing mutation, leaving the cell lines of (Figure 23) as the only confirmed cell line featuring missense mutations. As was mentioned in section 1.2.3, there is a wide distribution of thermodynamic disruption that can occur to a protein depending on the location and type of missense mutation, and there are certainly some mutations which could lead to effectively no change to the total ΔG of the protein and thus no change to the overall tertiary structure or stability of the protein, but could nonetheless impair function (Tokuriki et al. 2007). ASS1 (Figure 17) is a strange case where the control cells feature two bands, but the mutant cells only feature one, and while the mutation is unknown, we hypothesize the top band is the correct size for ASS1, as 38% biochemical activity in the patient was detected, however it is likely both bands for the control cell line represent ASS1, and the split in size is due to splicing or modification. A consistent clean band proved elusive for GAA (Figure 22), despite extensive attempts at optimization with two antibodies, significant dumbelling persisted. Clean bands for GAA in the literature have been few, but featured custom or no longer available antibodies (Moreland et al. 2005; Shimada et al. 2011). Despite this dumbelling issue, any significant increase in the mutant protein present would have been detected owing to the sensitivity of the technique. SGSH (Figure 23) contains two mutations previously confirmed with pulse-chase experiments to be degraded by the proteasome and upregulated by proteasome inhibitors as shown on western blots. This analysis was conducted using transgenic constructs of the mutant cDNA in baby hamster kidney cells; curiously no mention of similar analyses in patient cell lines has been reported (Muschol et al. 2004; Muschol et al. 2011). Furthermore unrelated groups observed residual sulphamidase activity of up to 3% (of normal controls) in cells with the

p.Ser298Pro mutation, therefore the catalytic function of the sulphamidase enzyme was believed to be negatively affected (Meyer et al. 2008). It seems likely that the mutation in the key functional domain results in normal levels of hypofunctioning protein in these patients. These same research groups have used transgenic over-expression to study these two missense mutations; they appear to be unstable when transgenically over-expressed in cell lines and also responsive to proteasome inhibitors (Muschol et al. 2004; Muschol et al. 2011).

4.2 Confounding Mechanisms

While this study attempts to focus on a simplistic model of proteasome degradation, whereby the transcription and translation of a missense mutant is identical to its wild-type cousin, with a shorter half-life due to its proteasomal degradation being the only difference, this does not appear to be the case. Similarly, the paradigm that bortezomib only interacts with the 26S proteasome and has no other consequences for the cell, besides an increase in ER stress, also does not appear to be the case.

There appears to be interplay between proteasome inhibition and mRNA levels, clearly shown in dysferlinopathy *in vitro* where a 24-fold increase in missense mutated protein levels was seen with a 4-fold increase in mRNA levels (Azakir et al. 2012). Dexamethasone, a steroid, in a similar study increased mRNA levels by 10-fold but dysferlin levels were only doubled. This shows that modulation to transcription and also to translation occurs in ways still unknown (Belanto et al. 2010).

Unstable mRNA is another confounding factor; mutations causing mRNA instability and thus reducing transcript would preclude effective bortezomib's treatment. This was explicitly mentioned by Coriell for one of the alleles for ADA (Figure 14).

Nonsense mediated decay is yet another consideration; a missense mutation may lead to the creation of an in-frame stop codon, then a premature termination codon (PTC) in the DNA, which results in nonsense mediated decay (NMD) for the mRNA transcript. There exists a process by which mRNA transcripts which have a premature stop codon are degraded in the nucleus, rather than transcribed possibly preventing the creation of small, dominant negative proteins harmful to the cell (Maquat 2004; Frischmeyer & Dietz 1999). This may have been a factor for ALDH3A2 (Figure 15) even though the transcript was nearly complete and did include the full functional domain.

Overexpression experiments have also added to the complexity, as outlined above they have shown mutant protein instability and bortezomib efficacy where none exists with normal patient cells. This is perhaps due to the extraordinarily high quantity of protein triggering mechanisms of protein ubiquitination and degradation which otherwise do not get triggered at endogenous levels. Specifically in the case of overexpressed unfolded mutant proteins, amyloids and aggregates are thought to occur. Mutant proteins are more likely to have hydrophobic residues exposed on their surface, relative to wild type proteins, making aggregates more thermodynamically favourable, and this could trigger the proteasome to respond in a way it would not endogenously (Gibson et al. 2013).

4.3 Stain-free Western Blotting

Since approximately 80-90% of all proteins are broken down by the 26S proteasome, it follows that the stain free proteome will all increase proportionally with bortezomib treatment masking any proportional upregulation of the mutated protein in question. The mutant protein must be upregulated to disproportionately high levels relative to the global proteome upregulation, in order for any change after protein normalization to be detected and for the treatment to be considered a success.

A significant improvement to the quantification and accuracy of the Western blotting performed in this study came from the transition to the Stain-free approach. This replaces the over-expressed and frequently outside-of-linear-range housekeeping proteins such as tubulin, GAPDH and actin with an in-gel chemistry; trihalocompound that upon UV exposure covalently reacts with a proportionally small amount of tryptophan residues of the proteins within the gel. This results in UV-induced fluorescence which offers comparative sensitivity to the Ponceau S molecule, however it can also be digitally quantified *in lieu* of a loading control, represents a larger proportion of the proteome and has a linear range up to 80 µg of protein loaded, twice the highest amount ever loaded for this study (Short & Posch 2011).

Further steps to ensure work was being done within the linear range of western blotting was performed by creating a serial dilution for every protein of interest, to ensure a reasonable amount of sample was loaded to stay within the linear range. Without such care there is a possibility that 40 µg of loaded sample may contain too much of a highly expressed protein of interest, masking any potential protein increase conferred by bortezomib treatment.

The more traditional film was used as the chemiluminescent detector instead of newer technologies, such as a charged-coupled device (CCD) digital camera. This was done primarily

for consistency with the equipment available at the beginning of the study. However film does have the advantage of a higher effective resolution. Rather than using relatively large discrete sensors, it has an effectively molecular resolution which is composed of a photochemical surface and also acquires signal for as long as is necessary. The disadvantage of film is the lower linear range, estimated to be 1-2 orders of magnitude, compared to the 2-5 orders of magnitude of a CCD sensor. Fortunately, to visualize a very publishable 400% increase in protein levels by bortezomib upregulation, would only require 0.6 orders of magnitude of linear range (log4) well within film's 1-2 orders of linear range (Bio Rad 2016).

4.4 Limitations

The greatest limitation incurred during this study was the limited selection in procuring cell lines from Coriell for a desired recessive disorder, and even when there were cells from one of the disorders of interest, missense mutated cell lines with reduced protein expression levels i.e. those cells which would in theory have the greatest potential to be upregulated by bortezomib were more often than not unavailable. To give this challenge some context, the following is an overview of the attrition rate in cell line identification. The approximately 2200 rare diseases published by Orphanet was fully examined; 400 were identified as being sufficiently high prevalence (>1 in 100 000), and also being monogenic (Orphanet 2016c). Approximately 100 of those were recessive disorders and had good expression levels in fibroblasts. Of those 100 disorders Coriell had cells for about 20, which included unknown genotypes and also premature stop codons as the only available cell lines. These approximately 20 are those tackled in this study, including lower quality (but still negative data) westerns which were not shown.

The next limitation was in acquiring a positive control cell line, to demonstrate bortezomib's efficacy during this study as found in the literature. Four groups were contacted, two did not respond, and the other two responded saying they no longer had any viable cells to send, as they had senesced.

The final limitation was in being able to choose and predict cell lines which would be ideal candidates to bortezomib. A database of protein expression levels for all missense mutants would have been extremely useful to focus efforts to those mutations that featured approximately 10% of wild type protein expression level which was met with success in the literature. Alternatively having an ability to accurately simulate the magnitude of protein misfolding *in silico* and thus being able to narrow down ideal candidates simply by knowing the missense mutation would have also been very useful. This was attempted with various protein folding simulation programs including: predictprotein (PredictProtein 2015), I-TASSER (R. Ambrish, K. Alper 2010) and finally YASARA (YASARA 2016) to no avail. The computational power required to accurately predict how a missense mutant would misfold in a cytoplasmic environment clearly lies in the future.

4.5 Conclusion

The amount of negative data acquired throughout this study suggests that bortezomib is not a generalizable treatment for recessive disorders that it might appear to be in the literature (Shimada et al. 2011; Azakir et al. 2012; Macías-Vidal et al. 2014). This is perhaps due to a publication bias whereby the presumably large number of cases of bortezomib being inefficacious goes unpublished, and the few times it is efficacious finds its way into the literature. Moreover, there are likely unknown molecular factors in effect, due to unexplainable phenomena such as bortezomib increasing mRNA levels during the treatment of dysferlinopathy (Azakir et al. 2012) and having greater success treating a splicing mutation for Pompe disease than treating its missense mutations (Shimada et al. 2011).

Additionally, there exist the confounding phenomena of over-expression experiments in the literature. It may be that in patient cells endogenous missense mutated proteins are at wild type levels and are unchanged by proteasomal inhibition, presumably exhibiting the disease phenotype by their mutation being in the functional domain. Then when over-expressed in non-human cell lines they are degraded relative to their wild type counterparts and thus modulated by proteasome inhibitor treatment. This certainly shows that independent of bortezomib and using proteasome inhibitors for recessive disorder treatment – there is a lot more molecular biology to uncover and understand.

4.6 Future Directions

While the future prospects of re-purposing bortezomib for recessive disorders appear bleak, the greater scope does not necessarily. Moving from a proteasome inhibitor which must be taken intravenously such as bortezomib to the recent FDA approval of ixazomib which is an oral proteasome inhibitor is an improvement in drug delivery and patient compliance (Takeda 2015; Muz et al. 2016). There is also the potential to find mutant cell lines of promising recessive disorders with demonstrated low protein levels by western blot, in the literature, and thus cooperate with these groups and acquire their cell lines. This might prove to be a more effective method than utilizing Coriell Cell Repositories and a more blind approach to mutant protein levels, providing such cell lines can be found in the literature and their respective research groups have available cells they are willing to share. Furthermore as knowledge of protein misfolding will increase; either with an aforementioned missense mutant protein expression database or powerful *in silico* tools backed up by increasing computing power, the ease of finding and testing that rare recessive disorder and mutation that is receptive to proteasome inhibitor treatment will greatly increase. Lastly, while gene and cell therapy ma

4.7 References

Adams, J. et al., 1999. Proteasome Inhibitors : A Novel Class of Potent and Effective Antitumor Agents Proteasome Inhibitors : A Novel Class of Potent and Effective Antitumor Agents. *Cancer Research*, 59(June 1), pp.2615–2622.

Almond, J.B. & Cohen, G.M., 2002. The proteasome: a novel target for cancer chemotherapy. *Leukemia*, 16(4), pp.433–43.

Azakir, B. a et al., 2012. Proteasomal inhibition restores biological function of mis-sense mutated dysferlin in patient-derived muscle cells. *The Journal of Biological Chemistry*, 287(13), pp.10344–54.

Belanto, J.J. et al., 2010. Dexamethasone induces dysferlin in myoblasts and enhances their myogenic differentiation. *Neuromuscular Disorders*, 20(2), pp.111–121.

Belgrader, P. & Maquat, L.E., 1994. Nonsense but not missense mutations can decrease the abundance of nuclear mRNA for the mouse major urinary protein, while both types of mutations can facilitate exon skipping. *Molecular and Cellular Biology*, 14(9), pp.6326–36.

Bio-Rad, 2016. What is alamarBlue. Available at: https://www.bio-rad-antibodies.com/alamarblue-cell-viability-assay-resazurin.html.

Bio Rad, 2016. Imaging Analysis and Documentation. Available at: http://www.bio-rad.com/en-ca/applications-technologies/imaging-analysis-documentation.

Blouin, J.-M. et al., 2013. Therapeutic potential of proteasome inhibitors in congenital erythropoietic porphyria. *Proceedings of the National Academy of Sciences of the United States of America*, 110(45), pp.18238–43.

Bolognese, A. et al., 2009. An NMR study of the bortezomib degradation under clinical use conditions. *Advances in Hematology*, 2009(Figure 2), pp.5–10.

Bross, P. et al., 1999. Protein misfolding and degradation in genetic diseases. *Human mutation*, 14(3), pp.186–98.

Bross, P. & Gregersen, N., 2003. *Protein Misfolding and Disease: Principles and Protocols*, Methods in Molecular Biology vol. 232: Totowa, N.J. : Humana Press, c2003.

Cartegni, L., Chew, S.L. & Krainer, A.R., 2002. Listening To Silence and Understanding Nonsense : Exonic Mutations That Affect Splicing. *Nature Reviews Genetics*, 3(April), pp.285–298.

Dill, K.A., 1990. Dominant Forces in Protein Folding. *Biochemistry*, 29(31), pp.7134–7155.

Dunker, A.K. et al., 2008. Function and structure of inherently disordered proteins. *Current Opinion in Structural Biology*, 18(6), pp.756–764.

EMBL-EBI, 2016. InterPro. Available at: http://www.ebi.ac.uk/interpro/sequencesearch/iprscan5-S20170123-203921-0185-17018732-pg [Accessed January 1, 2016].

EMEA, 2004. *Bortezomib Scientific Discussion*,

Fanin, M. & Angelini, C., 2015. Protein and genetic diagnosis of limb girdle muscular dystrophy type 2A: The yield and the pitfalls. *Muscle and Nerve*, 52(2), pp.163–173.

Frischmeyer, P. a & Dietz, H.C., 1999. Nonsense-mediated mRNA decay in health and disease. *Human molecular genetics*, 8(10), pp.1893–1900.

GeneCards, 2016. GeneCards. Available at: http://www.genecards.org/.

Gibson, T.J., Seiler, M. & Veitia, R. a, 2013. The transience of transient overexpression. *Nature methods*, 10(8), pp.715–21. Available at: http://dx.doi.org/10.1038/nmeth.2534.

Hartl, F., Bracher, Andreas., Hayer-Hartl, M., 2011. Molecular chaperones in protein folding and

proteostasis. *Nature*, 475, pp.324–332.
Hartl, F. & Hayer-Hartl, M., 2009. converging concepts of protein folding.pdf. *Nature Structural & Molecular Biology*, 16(6), pp.574–581.
Hirayama, S. et al., 2008. MKKS Is a Centrosome-shuttling Protein Degraded by Disease-causing Mutations via CHIP-mediated Ubiquitination. *Molecular Biology of the Cell*, 19(March), pp.899–911.
Janssen, 2015. VELCADE Product information. , 19(3), pp.143–145. Available at: www.janssen.com.au/files/Products/Velcade_PI.pdf.
Krawczak, M. et al., 2000. Human gene mutation database-a biomedical information and research resource. *Human mutation*, 15(1), pp.45–51.
Lavecchia, A. et al., 2013. Discovery of a Novel Small Molecule Inhibitor Targeting the Frataxin/ Ubiquitin Interaction via Structure-Based Virtual Screening and Bioassays. *Journal of Medicinal Chemistry*, 56, pp.2861–2873.
Lodish H, Berk A, Z.S., 2000. *Molecular Cell Biology 4th Edition*, New York. Available at: https://www.ncbi.nlm.nih.gov/books/NBK21726/.
Macías-Vidal, J. et al., 2014. The proteasome inhibitor bortezomib reduced cholesterol accumulation in fibroblasts from Niemann-Pick type C patients carrying missense mutations. *The FEBS journal*, 281, pp.4450–4466.
Maquat, L.E., 2004. Nonsense-Mediated mRNA Decay: Splicing, Translation And mRNP Dynamics. *Nature Reviews Molecular Cell Biology*, 5(2), pp.89–99.
Matyskiela, M.E., Lander, G.C. & Martin, A., 2013. Conformational switching of the 26S proteasome enables substrate degradation. *Nature Structural and Molecular Biology*, 20(7), pp.781–788.
Mcdowell, G.S. & Philpott, A., 2013. Non-canonical ubiquitylation: Mechanisms and consequences. *International Journal of Biochemistry and Cell Biology*, 45, pp.1833–1842.
Meyer, A. et al., 2008. The mutation p.Ser298Pro in the sulphamidase gene (SGSH) is associated with a slowly progressive clinical phenotype in mucopolysaccharidosis type IIIA (Sanfilippo A syndrome). *Human mutation*, 29(5), p.770.
Moreland, R.J. et al., 2005. Lysosomal acid a-glucosidase consists of four different peptides processed from a single chain precursor. *Journal of Biological Chemistry*, 280(8), pp.6780–6791.
Muschol, N. et al., 2011. Residual activity and proteasomal degradation of p.Ser298Pro sulfamidase identified in patients with a mild clinical phenotype of Sanfilippo A syndrome. *American Journal of Medical Genetics, Part A*, 155(June), pp.1634–1639.
Muschol, N. et al., 2004. Transport, enzymatic activity, and stability of mutant sulfamidase (SGSH) identified in patients with mucopolysaccharidosis type III A. *Human Mutation*, 23, pp.559–566.
Muz, B. et al., 2016. Spotlight on ixazomib: Potential in the treatment of multiple myeloma. *Drug Design, Development and Therapy*, 10, pp.217–226.
NIH, 2016. FAQs About Rare Diseases. *Genetic and Rare Diseases Information Center*. Available at: https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases.
Orphanet, 2016a. About Rare Diseases. Available at: http://www.orpha.net/consor/cgi-bin/Education_AboutRareDiseases.php?lng=EN.
Orphanet, 2015. Estimated prevalence of some rare diseases. *DG Health and Food Safety*. Available at: http://ec.europa.eu/health/rare_diseases/orphanet/report/index_en.htm.
Orphanet, 2016b. List of diseases or groups of diseases by decreasing prevalence. Available at:

http://www.orpha.net/orphacom/cahiers/docs/GB/Prevalence_of_rare_diseases_by_decreasing_prevalence_or_cases.pdf.
Orphanet, 2016c. Prevalence and incidence of rare diseases. *Orphanet*. Available at: http://www.orpha.net/orphacom/cahiers/docs/GB/Prevalence_of_rare_diseases_by_alphabetical_list.pdf.
Pekol, T. et al., 2005. Human metabolism of the proteasome inhibitor bortezomib: Identification of circulating metabolites. *Drug Metabolism and Disposition*, 33(6), pp.771–777.
PredictProtein, 2015. PredictProtein. Available at: https://www.predictprotein.org/.
R. Ambrish, K. Alper, Z.Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, 5(4), pp.725–738.
Rechsteiner, M. & Rogers, S.W., 1996. PEST sequences and regulation by proteolysis. *Trends in Genetics*, (July), pp.267–271.
SelleckChemicals, 2013. Selleck Chemicals. Bortezomib (PS-341). Available at: http://www.selleckchem.com/products/Bortezomib.html.
Sharma, M.C. & Goebel, H.H., 2005. Protein aggregate myopathies. *Neurology India*, 53(3), pp.273–9.
Shimada, Y. et al., 2011. Proteasome inhibitors improve the function of mutant lysosomal α-glucosidase in fibroblasts from Pompe disease patient carrying c.546G>T mutation. *Biochemical and biophysical research communications*, 415(2), pp.274–8.
Short, R. & Posch, A., 2011. Stain-Free Approach for Western Blotting. *Genetic Engineering & ...*, 31(20), pp.5–7.
Sorokin, a V, Kim, E.R. & Ovchinnikov, L.P., 2009. Proteasome system of protein degradation and processing. *Biochemistry*, 74(13), pp.1411–42.
Takeda, 2015. U.S. FDA Approves Takeda's NINLARO® (ixazomib), the First and Only Oral Proteasome Inhibitor to Treat Multiple Myeloma. Available at: https://www.takeda.com/news/2015/20151121_7185.html.
Tokuriki, N. et al., 2007. The Stability Effects of Protein Mutations Appear to be Universally Distributed. *Journal of Molecular Biology*, 369, pp.1318–1332.
Varshavsky, A., 2011. The N-end rule pathway and regulation by proteolysis. *Protein Science*, 20, pp.1298–1345.
Xu, G., Jiang, X. & Jaffrey, S.R., 2013. A mental retardation-linked nonsense mutation in cereblon is rescued by proteasome inhibition. *Journal of Biological Chemistry*, 288, pp.29573–29585.
YASARA, 2016. YASARA. Available at: http://yasara.org/.