Article

Computational investigation of growth hormone receptor Trp169Arg heterozygous mutation in a child with short stature†

Running head: *In-silico analysis of a GHR mutation*

William Farias Porto[1,2,3,a], Felipe Albuquerque Marques[1,4,5,a,b], Huri Brito Pogue[1], Maria Teresinha de Oliveira Cardoso[6,7]. Maria Gabriela Rodrigues do Vale[1], Állan da Silva Pires[1,2], Octavio Luiz Franco[1,2,8], Sérgio Amorim de Alencar[1,c]*, Robert Pogue[1,c]*

[1] Programa de Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília-DF, Brazil.
[2] Centro de Análises Proteômicas e Bioquímicas, Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília-DF, Brazil.
[3] Porto Reports, Brasília-DF, Brazil.
[4] Departamento de Farmácia, Universidade CEUMA, São-Luis-MA, Brazil
[5] Departamento de Biomedicina, Universidade CEUMA, São-Luis-MA, Brazil
[6] Curso de Medicina, Universidade Católica de Brasília, Brasília – DF, Brazil.
[7] Núcleo de Genética da Secretaria de Saúde do Distrito Federal, Brazil.
[8] S-Inova Biotech, Pós-graduação em Biotecnologia, Universidade Católica Dom Bosco, Campo Grande, MS, Brazil.
[a] These authors contributed equally to this work.
[b] Present address: Laboratório de Biotecnologia, Universidade CEUMA, São-Luis-MA, Brazil
[c] These authors contributed equally to this work.
*Corresponding authors: sergiodealencar@gmail.com (SAA) repogue@gmail.com (RP)
SGAN 916, Módulo B – Asa Norte, Brasília-DF, 70790-160, Brazil, Tel: +55-61-3448-7210, Fax: +55-61-3347-4797

Contract grant sponsor: Conselho Nacional de Desenvolvimento Científico e Tecnológico, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior and Fundação de Apoio a Pesquisa do Distrito Federal
Contract grant number: 564537/2010

Additional Supporting Information may be found in the online version of this article.

Received 16 January 2017; Revised 3 May 2017; Accepted 17 May 2017
Journal of Cellular Biochemistry

DOI 10.1002/jcb.26144

Abstract

Mutations in the growth hormone receptor (*GHR*) gene can cause disruption of the growth hormone signaling pathway, resulting in growth deficiency due to growth hormone (GH) resistance. Both recessive and apparently dominant mutations have been described in the literature. In order to shed some light on the molecular mechanism of partial growth hormone resistance caused by heterozygous mutations we performed an in-depth *in silico* analysis of a mutation found in a girl with a previous diagnosis of idiopathic short stature. An array of algorithms was used to predict pathogenicity and potential impact on the protein, and molecular modeling and molecular dynamic simulations were used to determine structural consequences. The results suggest that both of the possible single mutation-containing heteromeric GH-GHR complexes, as well as the double GHR mutant complex result in perturbation of complex structures, with altered ability of the GHR dimers to interact with the GH peptide. This article is protected by copyright. All rights reserved

Keywords: GHR; Short stature; *in silico* analysis; Docking; Molecular dynamics

1 INTRODUCTION

Growth Hormone (GH) is a peptide of 191 amino acids secreted by the anterior pituitary gland, and is important for linear growth. Its effects are mediated by the GH Receptor (GHR) (Godowski et al. 1989), present in many tissues such as muscle, liver and adipose (Mueller et al. 2012). This response to GH signaling is based on the high affinity between the GH and GHR proteins, which is due to hydrogen bonds and interactions between tryptophan atoms of the GHR homodimer, and residues at site 1 and site 2 of GH (Clackson et al. 1998). GH-GHR coupling triggers multiple signaling pathways, including the Janus Kinase/signal transducer and activator of transcription (JAK/STAT) pathways (Lanning & Carter-Su 2006), which stimulate the production of many signal transducers, among them insulin-like growth factor 1 (IGF-1). IGF-1 production is one of the most important consequences of GH-GHR interaction, responsible for clonal expansion and cellular maturation of chondrocytes in the growth plate (Cohen 2006). The interaction between GH and GHR involves a single GH molecule binding asymmetrically with a pre-formed GHR homodimer (Brown et al. 2005). Two different regions of GH, termed site 1 and site 2 interact with the separate monomers of GHR, with the same GHR residues largely being responsible for interaction with both sites (Wells 1996).

Mutations or deletions in the *GHR* gene result in lost (or decreased) ability to respond to GH, leading to different pathologies, including Laron Syndrome (LS; OMIM: 262500) (Laron et al. 1966), and short stature due to partial growth hormone insensitivity (OMIM: 604271). *GHR* is also found to be altered in a proportion of children initially diagnosed with idiopathic short stature (ISS). It is estimated that up to 5% of cases of ISS are caused by mutations in *GHR* (Goddard et al. 1997; Rakover et al. 2002).

Around 70 different mutations have been described for this gene, including missense, nonsense, deletions, frameshift and splice site mutations (David et al. 2011; Savage et al. 2006). Most of the reported mutations in *GHR* are homozygous or compound heterozygous (Aisenberg et al. 2010; Sobrier et al. 1997). However, Wood and colleagues observed that parents of children with LS had a marked short stature. As it is assumed that these individuals were heterozygous, it became clear that GHR heterozygous mutations with dominant negative effect could cause ISS (Woods et al. 1997). Subsequently, other mutations in this gene have been described with dominant negative effects (Ayling et al. 1997; Derr et al. 2011; Iida et al. 1998). However, these reports have not sought to clarify the mechanism by which some heterozygous mutations are pathogenic while others are apparently recessive and benign.

Recently, a wide number of *in silico* tools have been developed for assessing the impact of variants on the structure and function of the encoded protein (Hecht et al. 2013). This approach has proved particularly useful to assess the impact of point mutations that lead to amino acid residue changes in proteins, which are classified as deleterious or not by different prediction tools. Such tools have been applied to study several human proteins, including angiogenin (Padhi et al. 2013), anaplastic lymphoma kinase (Doss et al. 2014), aurora-A kinase (Kumar & Purohit 2014), aldosterone synthase (Jia et al. 2014), TP53 (Chitrala & Yeguvapalli 2014), guanylin peptides (Porto et al. 2015; Marcolino et al. 2016; Pires et al. 2017), α- and β-defensins (Porto, Nolasco, Pires, Fernandes, et al. 2016; Porto, Nolasco, Pires, Pereira, et al. 2016), phosphatase and tensin homolog (Khan et al. 2016), the low density lipoprotein receptor LA5 (Angarica et al. 2016), and fetal liver tyrosine kinase 3 (Swetha et al. 2016).

In summary, this study aims to use *in silico* approaches to examine the molecular defect in the GH signaling pathway responsible for short stature in a patient with a

heterozygous mutation in the *GHR* gene, through analysis of the structural impact caused by the alteration p.Trp169Arg.

2 MATERIALS AND METHODS

2.1 Patient

The legal representative of the patient signed an informed consent declaration allowing the use of DNA for research purposes. The research protocol was approved by the institutional Ethics Committee. The child (a girl) was born by vaginal delivery after an uneventful pregnancy. At birth she weighed 2810g (25th percentile), measured 45cm (3rd percentile), she had head circumference of 32cm (2nd-5th percentile) and Apgar score of 8-9. She held her head at 3 months, sat at 5 months, crawled at 6 months, and took her first steps and emitted her first words around 8 months of age. At 8 years of age the child had weight of 21.5kg (below P3), with height of 116.5 cm (also below P3). Cephalic perimeter was 51cm (P2-P50). There were no significant craniofacial deformities, though a slight posterior rotation of the ears was noted. She also had absent naso-labial philtrum, thick lips and signs of dental malocclusion. The feet were normal, as were her hands, apart from a shortening of the 5th metacarpal. All other physical exams were normal, as was her GH curve, while the basal level of IGF-1 was reduced. Her G-banded karyotype was unaltered (46,XX). The parents were of average height.

2.2 Molecular analyses

Genomic DNA was isolated from peripheral blood leukocytes using the Wizard Genomic DNA Purification Kit according to the manufacturer´s instructions (Promega®). The PCR reactions were performed in a Veriti® Thermal Cycler (Applied Biosystems®), using primers previously described to amplify all coding exons of *GHR* including the splice-junction regions (Bonioli et al. 2005). PCR products were sequenced on an ABI 3130 DNA

Sequencer (Applied Biosystems®). Electropherograms were analysed using CLC Main Workbench (CLC Bio®) with reference sequence of the *GHR* mRNA sequence (NM_000163.4) obtained from NCBI GenBank database.

2.3 Evolutionary Conservation Analysis

The ConSurf server is a tool for estimating the evolutionary conservation of amino acid positions in a protein molecule based on the phylogenetic relations between homologous sequences (Celniker et al. 2013). Using the growth hormone receptor isoform 1 precursor sequence (NP_000154), ConSurf, in ConSeq mode, a search was carried out for close homologous sequences using CSI-BLAST (3 iterations and 0.0001 of e-value cutoff) against the UNIREF-90 protein database (Suzek et al. 2007; Angermüller et al. 2012). The maximum number of homologs to collect was set at 150, and the minimal and maximal identity percentages between sequences were set as 35 and 95, respectively. The multiple sequence alignment and calculation methods were set at default (MAFFT-L-INS-i and Bayesian). The sequences were then clustered and highly similar sequences removed using CD-HIT (Li & Godzik 2006). Position-specific conservation scores were computed using the empirical Bayesian algorithm (Mayrose et al. 2004). The identification of exposed and buried side-chains was performed using a structure generated by HHPred (Söding et al. 2005).

2.4 *In silico* functional analyses of GHR point mutation

In order to evaluate the potential structural impact of the p.Trp169Arg point mutation, 16 *in silico* tools were used, divided into four different groups, as described by Porto and colleagues (Porto et al. 2015), with minor modifications. The p.Trp169Arg synthetic mutation was used as a negative control due to its deleterious character (Clackson et al. 1998).

2.4.1 Sequence homology-based methods

The following methods based on sequence homology principles were used to produce missense SNP functional predictions: Sorting Intolerant From Tolerant (SIFT) (Kumar et al. 2009), Provean (Choi et al. 2012), Mutation Assessor (Reva et al. 2011) and Panther (Mi et al. 2005).

2.4.2 Supervised learning methods

Supervised learning algorithms used for missense SNP impact prediction included neural networks (SNAP) (McLoone & Buchanan 2005), support vector machines (PhD-SNP; I-Mutant) (Capriotti et al. 2006; Capriotti et al. 2005) and SuSPect (Yates et al. 2014).

2.4.3 Protein sequence and structure-based methods

The following methods either combine information from protein sequence and structure or use protein structural information alone to analyze missense variants: PolyPhen [Adzhubei et al., 2010], mutation Cutoff Scanning Matrix (mCSM) (Pires et al. 2014), Fold-X (Guerois et al. 2002) and PoPMuSiC (Dehouck et al. 2011).

2.4.4 Consensus-based methods

In order to obtain a consensus score based on many different SNP impact prediction strategies, the following types of consensus software were used: Condel (González-Pérez & López-Bigas 2011), Meta-SNP (Capriotti et al. 2013), PON-P2 (Niroula et al. 2015) and PredictSNP (Bendl et al. 2014).

2.5 Molecular modeling

One hundred molecular models for each assembly (Trp/Trp, Trp/Arg, Arg/Trp, Arg/Arg and Ala/Ala) were constructed by comparative molecular modeling through MODELLER 9.14 (Webb & Sali 2014), using the structure of the receptor-hormone complex (PDB ID: 3HHR) (de Vos et al. 1992). The models were constructed using the default methods of automodel and environ classes from MODELLER. The final models were selected according to the discrete optimized protein energy score (DOPE score). This score assesses the energy of the model and indicates the best probable structures. The best models were evaluated through PROSA II (Wiederstein & Sippl 2007) and PROCHECK (Laskowski et al. 1993). PROCHECK checks the stereochemical quality of a protein structure through the Ramachandran plot, where good quality models are expected to have more than 90 % of amino acid residues in most favored and additional allowed regions, while PROSA II indicates the fold quality. Structure visualization was done in PyMOL (http://www.pymol.org).

2.6 Molecular Docking

Using the three-dimensional models as reference structures, the GH chain was removed from the complex and set as the ligand, while the B and C chains from the GHR were merged into a single chain and set as the receptor. The molecular docking was performed using the Hex 6.3 package (Ritchie & Venkatraman 2010), using shape plus electrostatics as the correlation type and energy minimization for post processing of each complex. The resulting complexes were clustered, using root mean square (RMS) cut-off of 3 Å. The cluster with highest affinity was selected as the preferable binding mode and the score of each complex in the cluster was averaged. The root mean square deviation (RMSD) of the

reference model and the resulting docking complex was calculated by using 3dSS server (Sumathi et al. 2006) by means of STAMP algorithm (Russell & Barton 1992).

2.7 Molecular Dynamics Simulation

The molecular dynamics simulations of the ensembles were carried out in a water environment, using the Single Point Charge water model (Berendsen et al. 1981). The analyses were performed by using the GROMOS96 43A1 force field and the GROMACS 4 computational package (Hess et al. 2008). The dynamics used the three-dimensional models as initial structures, immersed in water inside cubic boxes with a minimum distance of 0.8 nm between peptides and the edges of boxes. Sodium ions were also inserted at the complexes with negative charges in order to neutralize the system charge. Geometry of water molecules was constrained by using the SETTLE algorithm (Miyamoto & Kollman 1992). All atom bond lengths were linked by using the LINCS algorithm (Hess et al. 1997). Electrostatic corrections were made by Particle Mesh Ewald algorithm (Darden et al. 1993), with a cut-off radius of 1.4 nm in order to minimize the computational time. The same cut-off radius was also used for van der Waals interactions. The list of neighbors of each atom was updated every 20 simulation steps of 2 fs. The system underwent an energy minimization using 50,000 steps of the steepest descent algorithm. After that, the system temperature was normalized to 300 K for 100 ps, using the velocity rescaling thermostat (NVT ensemble). The system pressure was subsequently normalized to 1 bar for 100 ps, using the Parrinello-Rahman barostat (NPT ensemble). The systems with minimized energy, balanced temperature and pressure were simulated for 50 ns by using the leap-frog algorithm.

2.8 Analysis of Molecular Dynamics

The simulations were analyzed by means of the backbone root mean square deviation (RMSD), by means of g_rms utility from GROMACS. In addition, the interface area between the hormone and the receptor was calculated using equation 1: $Interface = \frac{A + B - AB}{2}$, where *A* denotes the solvent accessible surface area (SASA) of the hormone; *B*, the SASA of the receptor; and *AB*, the SASA of the complex. All SASA calculations were performed with the g_sas utility from GROMACS.

3 RESULTS

3.1 Sequencing

The sequencing of *GHR* showed a heterozygous substitution c.559T>C (NM_000163.4) (Figure 1), corresponding to the first nucleotide of codon 169 of the mature protein, in exon 6 of the *GHR* gene. The point mutation resulted in substitution of tryptophan by arginine (p.Trp169Arg) in the coded protein. One hundred population control chromosomes did not show this substitution. A search of online databases revealed that the variant has not been reported in dbSNP (www.ncbi.nlm.nih.gov/snp), Online Mendelian Inheritance in Man (OMIM: www.omim.org), Human Gene Mutation Data (HGMD: www.hgmd.cf.ac.uk), Exome Aggregation Consortium (ExAC: exac.broadinstitute.org) or the 1000 Genomes Database (www.1000genomes.org), nor was it encountered in the published literature. The variant was submitted to the ClinVar Database (ID: SUB1926751). Analysis of the rest of the coding sequence of *GHR* as well as splice boundaries did not reveal any further alterations. To assess the impact of this change on the GHR protein structure, a theoretical

analysis of the biochemistry of different amino acids was performed, followed by *in silico* modeling to assess its context within the protein and GH-GHR interaction.

3.2 Conservation analysis

Conservation analysis showed that the majority of tryptophan residues (Trp^{50}, Trp^{104}, Trp^{139}, Trp^{157}, Trp^{169} and Trp^{186}) have high levels of conservation, with grades from 8 to 9 (Figure 2). These data could be due to the fact that there is only one codon for tryptophan in the genetic code; however, the residues Trp^{16} and Trp^{76} were classified as extremely variable (Figure 2). In fact, Trp^{104} and Trp^{169} contribute substantially to the GH-GHR interaction, each accounting for >4.5 $kcal.mol^{-1}$ (Clackson et al. 1998) and both were predicted by Consurf as exposed and functional residues. The sequence reported here keeps the Trp^{104}, but Trp^{169} is altered to an arginine residue (p.Trp169Arg).

3.3 Prediction of deleterious effect of point mutation

According to a criteria established by Porto and colleagues (Porto et al. 2015), mutations classified as deleterious by at least three tools in each of the four groups of prediction tools are classified as convergent deleterious mutations. Using 16 different prediction tools to analyze the impact of p.Trp169Arg point mutation and our p.Trp169Ala control on GHR protein structure, only one tool (SuSPect) did not predict these amino acid residue changes as deleterious/destabilizing (Table 1). Hence, as the p.Trp169Arg convergent deleterious mutation is likely to impair protein structure and to have an impact on protein function, we further evaluated the structural impact caused by this mutation through docking analyses and molecular dynamics simulations.

3.4 Structural analyses

The p.Trp169Arg amino acid residue substitution is located in the binding site of GHR. Trp^{169} plays an important role in the stabilization of the GH-GHR interaction,

preventing electrostatic repulsion of two arginine residues, one from GH and the other from GHR. Due to the asymmetric character of GH, there are some differences in the interactions involving the residue Trp^{169} depending on the GHR chain. In this regard, since the mutation p.Trp169Arg is heterozygous, there are four possible GHR arrangements, named according to GH-site 1/2 interactions as Trp/Trp, Trp/Arg, Arg/Trp and Arg/Arg (Figure 3). The model assessments are in supplementary Table S1.

In order to evaluate the effect of this mutation on the complex, we performed docking analyses to verify the differences in the theoretical binding energies of hormone and receptor. The double wild-type was used as the positive control and the double p.Trp169Ala (Clackson et al. 1998) as the negative control. It was observed that the double mutant p.Trp169Arg and the heterodimer with p.Trp169Arg in chain 1 (which binds to GH site 1) showed decreased affinity for GH, as observed for our negative control, double p.Trp169Ala; however, the heterodimer with p.Trp169Arg in chain 2 showed virtually the same score of double wild-type (Figure 3). These data indicated that once chain 1 is wild-type, it does not matter if chain 2 is mutated.

As the docking experiments only take into account the static structures, we performed a short molecular dynamics simulation (50 ns) in order to validate the docking results. The RMSD evolution along the simulations (Figure 4A) indicated that the complexes with at least one chain carrying the mutation suffer a slight structural modification when compared to the double wild type, which also occurs with the double Trp169Ala complex. Therefore, despite the complex Trp/Arg seeming to have the same affinity of Trp/Trp to GH, it undergoes a structural modification that could lead to malfunction.

Concerning the interface area of the complex, only the complex with Arg/Trp showed a larger interface area, similar to the negative control, Ala/Ala; and the double mutant Arg/Arg showed a reduction in the interface, indicating a repulsive character; while the

complex with Trp/Arg showed a similar behavior to the double wild type. However, according to RMSD, as Trp/Arg has a structural modification, even with a similar interface, it could be unable to activate the JAK2 molecules and unleash the signaling cascade as described by Brown and co-workers (Brown et al. 2005).

4 DISCUSSION

Currently, 80% of short stature cases initially diagnosed as idiopathic do not have their mechanisms elucidated. It is known that up to 15% of cases are related to mutations in the Short Stature Homeobox gene (*SHOX*) (Binder 2011; Huber et al. 2006) and 5% are due to mutations in the *GHR* (Attie 2000; Bonioli et al. 2005). The patient described here had short stature apparently caused by a mutation in the *GHR* gene, and mutations in *SHOX* were not detected (data not shown). Thus, the heterozygous p.Trp169Arg mutation seems to cause partial growth hormone insensitivity, affecting a key amino acid in the GHR extracellular domain. As 80% of the cases have unknown mechanisms, we have to be cautious about our assumptions regarding the p.Trp169Arg mutation. Due to the complex molecular genetic mechanisms involved in linear growth, it is possible that other polimorphisms in the genome are also influencing the patient´s final height. However, according to Attie (2000), we can discard mutations in some genes in the same pathway, as they usually result in a severe phenotype, different from the patient described here.

We found a heterozygous substitution of an extremely conserved tryptophan (Figure 1 and 2) for arginine at residue 169 of the GHR protein, which was predicted as convergent deleterious (Table 1). The main contributions to the GH-GHR interaction are from hydrophobic side-chains of GHR, with two tryptophan (Trp^{104} and Trp^{169}) side-chains each accounting for >4.5 $kcal.mol^{-1}$ (Clackson et al. 1998). Together, these amino acid residues are responsible for 20% of the total contact area between GH and GHR, mainly due to the indole rings (Clackson et al. 1998). Tryptophan at position 169 is surrounded by a hydrophylic

periphery, formed by arginine residues. In the complex, Trp^{169} is located between Arg^{43} from GHR and Arg^{64} (site 1) or Arg^{16} (site 2) of GH and it is able to interact with Ile^{179} in site 1 of GH. According to Clackson and co-workers (Clackson et al. 1998), the monomeric receptor with the mutation p.Trp169Ala is unable to interact with GH *in vitro*, perhaps due to the repulsion between the arginine residues and the lack of interaction with Ile^{179}. The substitution found in the current patient, p.Trp169Arg, results in three arginine residues side by side in the complex, magnifying the repulsion and modifying the angles and tertiary structure of the GH-GHR interaction domain. Besides, by switching Trp to Arg there is a possible absence of interaction between Arg and Ile residues. However, according to (Goddard et al. 1997), the monomeric receptor with the mutation p.Glu44Lys, which is also in the interaction site, is able to interact with GH *in vitro*, but with reduced affinity.

The fact that, *in vivo*, GHR is a homodimer and performs an asymmetric interaction with GH, together with the fact that the current patient is heterozygous for p.Trp169Arg, add a new layer of complexity to the analysis of the effect of such mutation. According to Brown et al. (2005), the signal transduction by GH-GHR interaction occurs in four steps, (i) GHR dimer is formed; (ii) GH site 1 binds to GHR chain 1; (iii) GHR chain 2 binds to site 2 and (iv) there occurs a conformational change in the complex, involving a rotation and a vertical movement. This movement generated a rotation in the transmembrane domains that activates the JAK2 by transphosphorylation (Brown et al. 2005).

Our *in silico* approach covers the steps after the GHR dimer formation, as the mutation p.Trp169Arg is not in the dimer interface. First, by docking analysis, we observed that chain 1 must be wild-type (Figure 3), therefore, assemblies Arg/Trp and Arg/Arg are unable to perform the first interaction with GH. Second, we performed molecular dynamics simulations to analyze the binding of chain 2 to GH and then the conformational change.

The simulation indicated the complex maintenance during 50 ns, it should be observed that there is a degree of bias in the simulations, since they started with the hormone already bonded to the receptor, and so are based on the presumption that the complex can form in the presence of at least one mutated monomer, which may not actually be the case, especially for Arg/Trp and Arg/Arg complexes. Furthermore, due to the box size and the time of the simulation, it was not possible to observe the separation of the complex.

The simulations confirmed the docking results for the Arg/Trp and Arg/Arg complexes. We observed a reduction in the interface area of the double mutant complex (Figure 4B), and Arg/Trp showed the same behavior as our negative control (double p.Trp169Ala), which could indicate an inability of GH to bind to Arg/Trp or Arg/Arg receptors. In contrast, the heterodimeric receptor Trp/Arg seems to be able to interact with GH site 2, as the interface of interaction was virtually equal to the double wild-type (Figure 4B). Nevertheless, according to the RMSD evolution during the simulation, all of the mutant complexes underwent structural modifications differing from the double wild-type (Figure 4A) and, due to the structural modification, it may be that receptor with the mutation are inactive, resulting in only 25% of receptors interacting adequately with GH.

It is known that in cases where only 25% of the receptors are functional, patients seem to produce lower amounts of IGF-1, as a response to GH stimulus (Iida et al. 1998). These are generally presumed to have a dominant-negative effect on the receptor. Besides, Zhou et al. (Zhou et al. 1997) demonstrated that for the heterozygous GHR knockout mouse, the growth is mildly retarded, while for double knockout this condition is severe; and maybe, a similar scenario could be assumed for the patient described here, mainly due to the fact that no mutations in SHOX were observed. Nevertheless this is the first study that examines separately the effects of mutation in the two possible heterodimers. Since the published

literature provides evidence of heterozygous *GHR* mutations causing growth and skeletal defects, our data provide insight into the possible molecular mechanisms involved.

While the present data offer an important insight into the primary molecular defect caused by the p.Trp169Arg substitution in GHR, it is important to highlight that mutations in other positions of GHR could affect other steps of signal transduction. Our data describe the interaction in the complex, however, other mutations could disrupt the formation of GHR dimer, as the case of p.Asp152His (Duquesnoy et al. 1994).

The question remains as to whether some other aspects of the heterozygous dimers reduce their functionality, such as reduced capacity to form the GH-GHR complex, or to transduce signals. The mutation described here may disrupt the interaction between GH and GHR. As such, if both alleles of GHR were expressed equally, one would expect to observe up to 75% of receptors with malfunction. We could not discard other factors, as in 80% of similar cases the mechanism that causes short stature is unknown; also the parental DNA would be valuable to determine the status of the heterozygous mutation. We expect that, in a near future, our *in silico* approach could be used to study other mutations on GHR, as a part of personalized medicine programs.

5 ACKNOWLEDGEMENTS

We thank the patient and her legal representative for agreeing to this study. This work was funded by the Rede Pro-Centro-Oeste. We also thank Dr. Elisa Lamback for interpreting the GH curve. The authors declare that there is no conflict of interest associated with this research.

6 REFERENCES

Aisenberg, J. et al., 2010. Atypical GH insensitivity syndrome and severe insulin-like growth factor-I deficiency resulting from compound heterozygous mutations of the GH receptor, including a novel frameshift mutation affecting the intracellular domain. *Hormone research in paediatrics*, 74(6), pp.406–11. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20606392.

Angarica, V.E., Orozco, M. & Sancho, J., 2016. Exploring the complete mutational space of the LDL receptor LA5 domain using molecular dynamics: linking SNPs with disease phenotypes in familial hypercholesterolemia. *Human molecular genetics*, 25(6), pp.1233–46. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26755827.

Angermüller, C., Biegert, A. & Söding, J., 2012. Discriminative modelling of context-specific amino acid substitution probabilities. *Bioinformatics (Oxford, England)*, 28(24), pp.3240–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23080114 [Accessed June 8, 2015].

Attie, K.M., 2000. Genetic studies in idiopathic short stature. *Current opinion in pediatrics*, 12(4), pp.400–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10943824.

Ayling, R.M. et al., 1997. A dominant-negative mutation of the growth hormone receptor causes familial short stature. *Nature genetics*, 16(1), pp.13–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9140387.

Bendl, J. et al., 2014. PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. P. P. Gardner, ed. *PLoS computational biology*, 10(1), p.e1003440.

Berendsen, H. et al., 1981. Interaction models for water in relation to protein hydration. In B. Pullman, ed. *Intermolecular Forces*. Springer, pp. 331–342.

Binder, G., 2011. Short stature due to SHOX deficiency: genotype, phenotype, and therapy.

Hormone research in paediatrics, 75(2), pp.81–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21325865.

Bonioli, E. et al., 2005. Heterozygous mutations of growth hormone receptor gene in children with idiopathic short stature. *Growth hormone & IGF research : official journal of the Growth Hormone Research Society and the International IGF Research Society*, 15(6), pp.405–10. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16213173.

Brown, R.J. et al., 2005. Model for growth hormone receptor activation based on subunit rotation within a receptor dimer. *Nature structural & molecular biology*, 12(9), pp.814–21. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16116438.

Capriotti, E. et al., 2005. Predicting protein stability changes from sequences using support vector machines. *Bioinformatics (Oxford, England)*, 21 Suppl 2, p.ii54-8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16204125.

Capriotti, E., Altman, R.B. & Bromberg, Y., 2013. Collective judgment predicts disease-associated single nucleotide variants. *BMC genomics*, 14 Suppl 3, p.S2.

Capriotti, E., Calabrese, R. & Casadio, R., 2006. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics (Oxford, England)*, 22(22), pp.2729–34.

Celniker, G. et al., 2013. ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. *Israel Journal of Chemistry*, 53(3–4), pp.199–206.

Chitrala, K.N. & Yeguvapalli, S., 2014. Computational screening and molecular dynamic simulation of breast cancer associated deleterious non-synonymous single nucleotide polymorphisms in TP53 gene. *PloS one*, 9(8), p.e104242.

Choi, Y. et al., 2012. Predicting the functional effect of amino acid substitutions and indels. A. G. de Brevern, ed. *PloS one*, 7(10), p.e46688.

Clackson, T. et al., 1998. Structural and functional analysis of the 1:1 growth

hormone:receptor complex reveals the molecular basis for receptor affinity. *Journal of molecular biology*, 277(5), pp.1111–28. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9571026.

Cohen, P., 2006. Overview of the IGF-I system. *Hormone research*, 65 Suppl 1, pp.3–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16508327.

Darden, T., York, D. & Pedersen, L., 1993. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12), p.10089.

David, A. et al., 2011. Evidence for a continuum of genetic, phenotypic, and biochemical abnormalities in children with growth hormone insensitivity. *Endocrine reviews*, 32(4), pp.472–97. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21525302.

Dehouck, Y. et al., 2011. PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC bioinformatics*, 12, p.151.

Derr, M.A. et al., 2011. The growth hormone receptor (GHR) c.899dupC mutation functions as a dominant negative: insights into the pathophysiology of intracellular GHR defects. *The Journal of clinical endocrinology and metabolism*, 96(11), pp.E1896-904. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21900382.

Doss, C.G.P. et al., 2014. Integrating in silico prediction methods, molecular docking, and molecular dynamics simulation to predict the impact of ALK missense mutations in structural perspective. *BioMed research international*, 2014, p.895831. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25054154.

Duquesnoy, P. et al., 1994. A single amino acid substitution in the exoplasmic domain of the human growth hormone (GH) receptor confers familial GH resistance (Laron syndrome) with positive GH-binding activity by abolishing receptor homodimerization. *The EMBO journal*, 13(6), pp.1386–95. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8137822.

Goddard, A.D. et al., 1997. Partial growth-hormone insensitivity: the role of growth-hormone receptor mutations in idiopathic short stature. *The Journal of pediatrics*, 131(1 Pt 2), pp.S51-5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9255229.

Godowski, P.J. et al., 1989. Characterization of the human growth hormone receptor gene and demonstration of a partial gene deletion in two patients with Laron-type dwarfism. *Proceedings of the National Academy of Sciences of the United States of America*, 86(20), pp.8083–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2813379.

González-Pérez, A. & López-Bigas, N., 2011. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *American journal of human genetics*, 88(4), pp.440–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3071923&tool=pmcentrez&rendertype=abstract [Accessed February 13, 2016].

Guerois, R., Nielsen, J.E. & Serrano, L., 2002. Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *Journal of Molecular Biology*, 320(2), pp.369–387.

Hecht, M., Bromberg, Y. & Rost, B., 2013. News from the protein mutability landscape. *Journal of molecular biology*, 425(21), pp.3937–48. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23896297.

Hess, B. et al., 2008. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(3), pp.435–447.

Hess, B. et al., 1997. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12), pp.1463–1472.

Huber, C. et al., 2006. High incidence of SHOX anomalies in individuals with short stature. *Journal of medical genetics*, 43(9), pp.735–9. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/16597678.

Iida, K. et al., 1998. Growth hormone (GH) insensitivity syndrome with high serum GH-binding protein levels caused by a heterozygous splice site mutation of the GH receptor gene producing a lack of intracellular domain. *The Journal of clinical endocrinology and metabolism*, 83(2), pp.531–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9467570.

Jia, M. et al., 2014. Computational analysis of functional single nucleotide polymorphisms associated with the CYP11B2 gene. *PloS one*, 9(8), p.e104311. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4125216&tool=pmcentrez&rendertype=abstract.

Khan, I. et al., 2016. Prediction of functionally significant single nucleotide polymorphisms [SNPs] in PTEN tumor suppressor gene: An in silico approach. *Biotechnology and Applied Biochemistry*, p.n/a-n/a. Available at: http://doi.wiley.com/10.1002/bab.1483.

Kumar, A. & Purohit, R., 2014. Use of Long Term Molecular Dynamics Simulation in Predicting Cancer Associated SNPs A. D. MacKerell, ed. *PLoS Computational Biology*, 10(4), p.e1003318.

Kumar, P., Henikoff, S. & Ng, P.C., 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols*, 4(7), pp.1073–81.

Lanning, N.J. & Carter-Su, C., 2006. Recent advances in growth hormone signaling. *Reviews in endocrine & metabolic disorders*, 7(4), pp.225–35. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17308965.

Laron, Z., Pertzelan, A. & Mannheimer, S., 1966. Genetic pituitary dwarfism with high serum concentation of growth hormone--a new inborn error of metabolism? *Israel journal of medical sciences*, 2(2), pp.152–5. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/5916640.

Laskowski, R. et al., 1993. {PROCHECK}: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 26, pp.283–291.

Li, W. & Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)*, 22(13), pp.1658–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16731699 [Accessed July 10, 2014].

Marcolino, A.C.S. et al., 2016. Structural impact analysis of missense SNPs present in the uroguanylin gene by long-term molecular dynamics simulations. *Journal of theoretical biology*, 410, pp.9–17. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27620667.

Mayrose, I. et al., 2004. Comparison of site-specific rate-inference methods for protein sequences: Empirical Bayesian methods are superior. *Molecular Biology and Evolution*, 21(9), pp.1781–1791.

McLoone, E.M. & Buchanan, T.A.S., 2005. Unusual macular lesions in a patient with neurofibromatosis type-1. *International ophthalmology*, 26(3), pp.115–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16957876.

Mi, H. et al., 2005. The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic acids research*, 33(Database issue), pp.D284-8.

Miyamoto, S. & Kollman, P.A., 1992. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry*, 13(8), pp.952–962.

Mueller, K.M. et al., 2012. Hepatic growth hormone and glucocorticoid receptor signaling in body growth, steatosis and metabolic liver cancer development. *Molecular and cellular endocrinology*, 361(1–2), pp.1–11. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22564914.

Niroula, A., Urolagin, S. & Vihinen, M., 2015. PON-P2: Prediction Method for Fast and

Reliable Identification of Harmful Variants. *PloS one*, 10(2), p.e0117380.

Padhi, A.K., Jayaram, B. & Gomes, J., 2013. Prediction of functional loss of human angiogenin mutants associated with ALS by molecular dynamics simulations. *Scientific reports*, 3, p.1225.

Pires, Á.S. et al., 2017. Theoretical structural characterization of lymphoguanylin: A potential candidate for the development of drugs to treat gastrointestinal disorders. *Journal of theoretical biology*, 419, pp.193–200. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28214543.

Pires, D.E. V, Ascher, D.B. & Blundell, T.L., 2014. mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics (Oxford, England)*, 30(3), pp.335–42. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24281696.

Porto, W.F., Nolasco, D.O., Pires, Á.S., Fernandes, G.R., et al., 2016. HD5 and HBD1 variants' solvation potential energy correlates with their antibacterial activity against Escherichia coli. *Biopolymers*, 106(1), pp.43–50. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26567133 [Accessed November 15, 2015].

Porto, W.F., Nolasco, D.O., Pires, Á.S., Pereira, R.W., et al., 2016. Prediction of the impact of coding missense and nonsense single nucleotide polymorphisms on HD5 and HBD1 antibacterial activity against Escherichia coli. *Biopolymers*, 106(5), pp.633–644. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27160989 [Accessed November 6, 2016].

Porto, W.F., Franco, O.L. & Alencar, S.A., 2015. Computational analyses and prediction of guanylin deleterious SNPs. *Peptides*, 69, pp.92–102. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25899674 [Accessed April 23, 2015].

Rakover, Y. et al., 2002. Can exaggerated response to a GH provocative test identify patients with partial GH insensitivity syndrome? *European journal of endocrinology*, 146(3),

pp.319–23. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11888837.

Reva, B., Antipin, Y. & Sander, C., 2011. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic acids research*, 39(17), p.e118. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3177186&tool=pmcentrez&rendertype=abstract [Accessed July 15, 2014].

Ritchie, D.W. & Venkatraman, V., 2010. Ultra-fast FFT protein docking on graphics processors. *Bioinformatics (Oxford, England)*, 26(19), pp.2398–405. Available at: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btq444 [Accessed March 31, 2017].

Russell, R.B. & Barton, G.J., 1992. Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins*, 14(2), pp.309–23. Available at: http://doi.wiley.com/10.1002/prot.340140216 [Accessed March 31, 2017].

Savage, M.O. et al., 2006. Endocrine assessment, molecular characterization and treatment of growth hormone insensitivity disorders. *Nature clinical practice. Endocrinology & metabolism*, 2(7), pp.395–407. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16932322.

Sobrier, M.L. et al., 1997. Nine novel growth hormone receptor gene mutations in patients with Laron syndrome. *The Journal of clinical endocrinology and metabolism*, 82(2), pp.435–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9024232.

Söding, J., Biegert, A. & Lupas, A.N., 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic acids research*, 33(Web Server issue), pp.W244-8. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1160169&tool=pmcentrez&r

endertype=abstract [Accessed July 9, 2014].

Sumathi, K. et al., 2006. 3dSS: 3D structural superposition. *Nucleic acids research*, 34(Web Server issue), pp.W128-32. Available at: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkl036 [Accessed March 31, 2017].

Suzek, B.E. et al., 2007. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics (Oxford, England)*, 23(10), pp.1282–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17379688 [Accessed February 9, 2015].

Swetha, R.G., Ramaiah, S. & Anbarasu, A., 2016. Molecular Dynamics Studies on D835N Mutation in FLT3-Its Impact on FLT3 Protein Structure. *Journal of cellular biochemistry*, 117(6), pp.1439–45. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26566084.

de Vos, A.M., Ultsch, M. & Kossiakoff, A.A., 1992. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science (New York, N.Y.)*, 255(5042), pp.306–12. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1549776.

Webb, B. & Sali, A., 2014. Comparative Protein Structure Modeling Using MODELLER. *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]*, 47, p.5.6.1-5.6.32. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25199792 [Accessed November 7, 2014].

Wells, J.A., 1996. Binding in the growth hormone receptor complex. *Proceedings of the National Academy of Sciences of the United States of America*, 93(1), pp.1–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8552582.

Wiederstein, M. & Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research*, 35(Web Server issue), pp.W407-10. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1933241&tool=pmcentrez&rendertype=abstract [Accessed September 8, 2014].

Woods, K.A. et al., 1997. Phenotype: genotype relationships in growth hormone insensitivity syndrome. *The Journal of clinical endocrinology and metabolism*, 82(11), pp.3529–35. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9360502.

Yates, C.M. et al., 2014. SuSPect: enhanced prediction of single amino acid variant (SAV) phenotype using network features. *Journal of molecular biology*, 426(14), pp.2692–701.

Zhou, Y. et al., 1997. A mammalian model for Laron syndrome produced by targeted disruption of the mouse growth hormone receptor/binding protein gene (the Laron mouse). *Proceedings of the National Academy of Sciences of the United States of America*, 94(24), pp.13215–20. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9371826 [Accessed March 31, 2017].

7 FIGURE LEGENDS

Figure 1 - Heterozygous point mutation in GHR gene. The reference sequence is on top, corresponding to Exon 6 deposited in NCBI (NG_011688.1). All mutations are highlighted in red. The patient electropherogram shows two substitutions in heterozygous: a benign SNP (rs6179) at position 558 and mutation at position 559 of the mRNA coding sequence (Thymine to Cytosine), which corresponds to the first nucleotide of codon 169, resulting in a substitution of an Arginine by Tryptophan (p.Trp169Arg). The control electropherogram shows the benign SNP rs6179 in homozygous at position 558.

Figure 2 – Evolutionary conservation of GHR amino acid residues obtained from multiple sequence alignment by ConSurf. The amino acids colors are based on their conservation grades and conservation levels. A grade of 1 indicates rapidly evolving (variable) sites, grade of 5 indicates sites that are evolving at an average rate, and 9 indicates slowly evolving (evolutionarily conserved) sites. The colors are explained in the figure. e – represents an exposed residue according to the neural-network algorithm; b - a buried residue according to the neural-network algorithm; f - a predicted functional residue (highly conserved and exposed); s - a predicted structural residue (highly conserved and buried).

Figure 3 – Docking analyses of the four complex arrangements. The GH is represented in gray, while the chains of GHR are in blue (wild type), purple (Trp169Arg) or red (Trp169Ala). The hex score was obtained by averaging the scores for the best docking cluster and the RMSD is relative to their respective comparative molecular model.

Figure 4 – Molecular dynamics analyses. (A) The backbone's RMSD evolution along the simulation time, showing a structural difference from the double wild type receptor (Trp/Trp) starting from 25 ns of simulation. (B) The interface area of the GH-GHR complex.

Table 1 – Prediction results of *GHR* p.W169R point mutation analyzed by 16 prediction tools classified in four different groups.

		Sequence-based[b]				SLM-based[b]				Consensus-based[b]				Structure-based[b]			
Point Mutation	**Amino Acid Change[a]**	**SIFT**	**Provean**	**Mutation Assessor**	**Panther**	**PhD-SNP**	**I-Mutant**	**SNAP**	**SuSPect**	**Condel**	**MetaSNP**	**PON-P2 Predict**	**SNP**	**PolyPhen**	**mCSM**	**Fold-X**	**PoP MuSiC**
-	p.W169A	D	D	D	D	D	D	D	N	D	D	D	D	D	DT	DT	D
c.558T>C	p.W169R	D	D	D	D	D	D	D	N	D	D	D	D	D	DT	DT	D

[a] GHR amino acid position is relative to the mature protein, PDB Accession number 3HHR. The p.W169A mutation is synthetic with known deleterious character.
[b] N: Neutral; D: Deleterious; ST: stabilizing; DT: Destabilizing; U: Unknown.

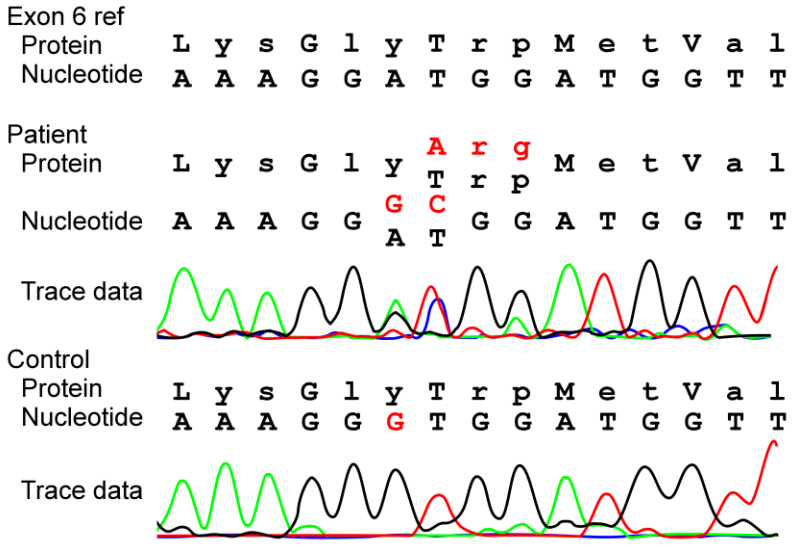

Figure 1

Figure 2

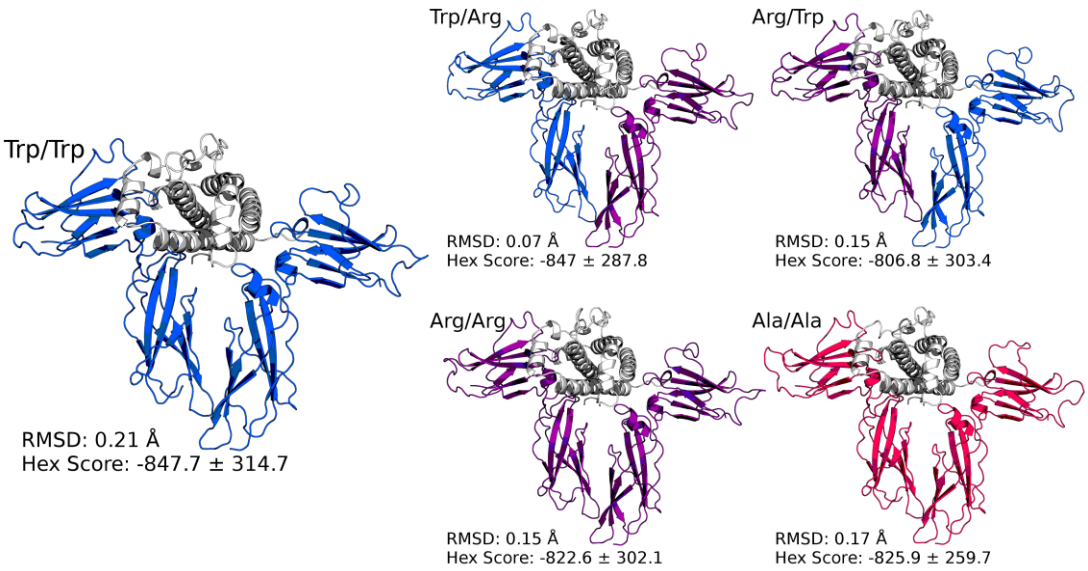

Figure 3

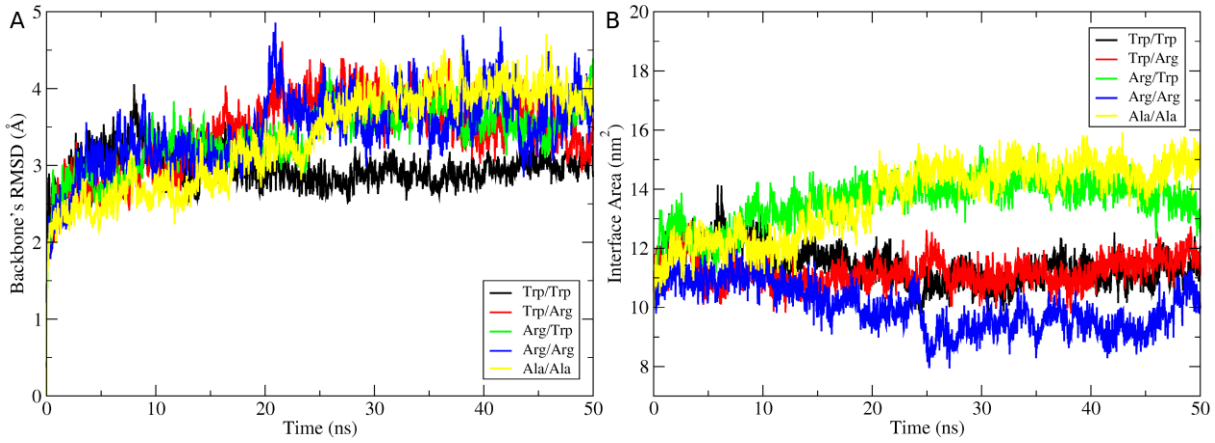

Figure 4