SNPeffect overview

September 6, 2014

Chapter 1

Introduction

In this analysis, we focus on the effect of mutation on aggregation propensity (TANGO), amyloid propensity (WALTZ) and chaperone binding (LIMBO). It is important to point out that beta aggregation is mediated by short stretches that need to become exposed by (partial) unfolding before they can actually nucleate protein aggregation. Therefore, when analysing mutations, we consider two important effects.

- the intrinsic aggregation propensity, and
- the stability of the protein.

As protein folding involves a large proportion of amino acid residues, most mutations have their effect by acting on stability, i.e. by exposing aggregation prone regions. In order to estimate the likelihood that a given short stretch may become exposed, we employ FoldX to caclulate the effect on the structural stability.
So, the presence of TANGO, WALTZ or LIMBO regions does not necessarily implies that the protein readily forms aggregates, amyloid or exposes chaperone binding regions respectively. Such regions are normally buried in the protein core, but when these regions become exposed due to other factors (e.g. by a structurally destabilizing mutation), it can become more prominent.

Chapter 2

Phenotypic analysis of R231G in uniprot id p08294

2.1 TANGO prediction

TANGO predicts the aggregation prone regions in a protein sequence. The total TANGO score for you protein is 641.26. Mutations can increase (dTANGO >50), decrease (dTANGO <-50) or not affect aggregation propensity (dTANGO between -50 and 50). For this mutation, dTANGO equals -1.85 which means that the mutation does not affect the aggregation tendency of your protein.

In figure 2.1 and 2.2 the position of the TANGO stretches in the wild type and variant

Figure 2.1: Bar representation of the TANGO windows present in the wild type (top) and mutant protein (bottom). In the bar representation, the position of the aggregating stretches is visualised in red, and the dashed vertical line in the variant indicates the position of the variant residue.

Table 2.1: TANGO regions in variant and wild type. For each TANGO region, the start, end, sequence and score is given.

Number	Start	End	Stretch	Score
Wild Type				
1	1	13	LALLCSCLLLAA	22.86
2	77	84	VTGVVLF	40.15
Mutant				
1	1	13	LALLCSCLLLAA	22.79
2	77	84	VTGVVLF	40.06

protein are visualized, represented by respectively a bar or profile representation. In table 2.1, the short stretches are listed for both wild type and mutant. To compare the effect of the mutation to the WT, we also show a Difference profile (Figure 2.3), that plots the difference

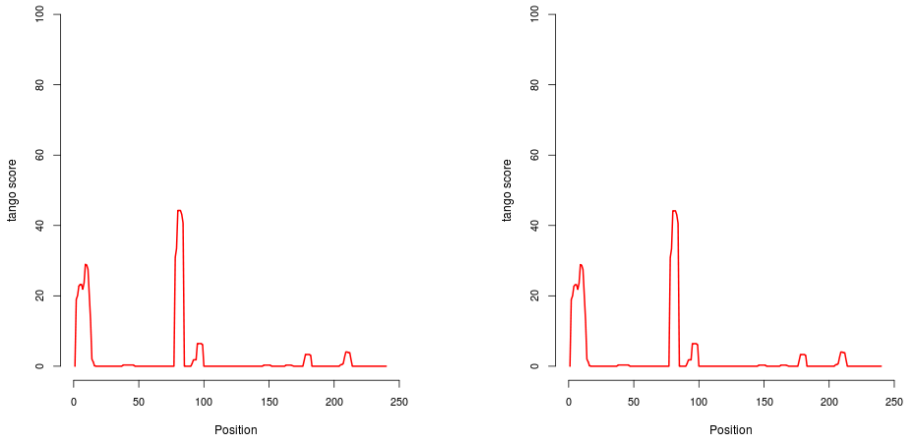

Figure 2.2: Profile representation of the TANGO stretches in the wild type (left) and mutant (right) protein. This graph plots the per-residue TANGO aggregation score of the wild type and variant protein. From left to right, all residue scores from the N-terminus to the C-terminus are plotted.

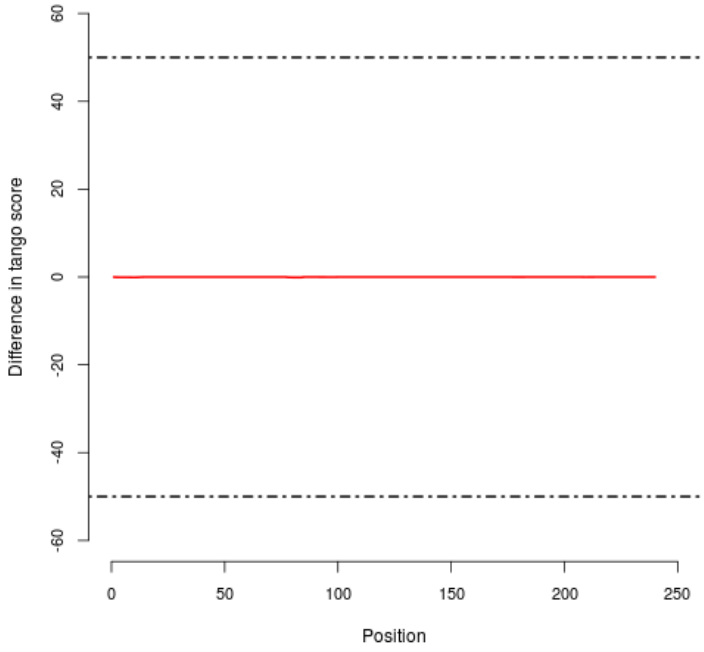

Figure 2.3: Difference in TANGO aggregation between WT and variant. This graph plots the per-residue TANGO aggregation score diference between WT protein and the variant. From left to right, all TANGO score differences from the N-terminus to the C-terminus are plotted. A flat line indicates that the variant does not alter the aggregation profile of the protein. Positive peaks indicate increased aggregation tendency due to this mutation. Negative peaks indicate decreased aggregation tendency due to this mutation.

between WT protein and the variant.

2.2 WALTZ prediction

WALTZ is an algorithm that accurately and specifically predicts amyloid-forming regions in protein sequences. It is thus more specific in terms of aggregate morphology than TANGO. The total WALTZ score for your protein is 165.98 and mutations can increase (dWALTZ >50), decrease (dWALTZ <-50) or not affect amyloid propensity (dWALTZ between -50 and 50). In this case, dWALTZ equals 0.16 which means that the mutation does not affect the amyloid propensity of your protein.

In figure 2.4 and 2.5 the position of the WALTZ stretches in the wild type and variant

Figure 2.4: Bar representation of the WALTZ windows present in the wild type (top) and mutant protein (bottom). In the bar representation, the position of the aggregating stretches is visualised in blue, and the dashed vertical line in the variant indicates the position of the variant residue.

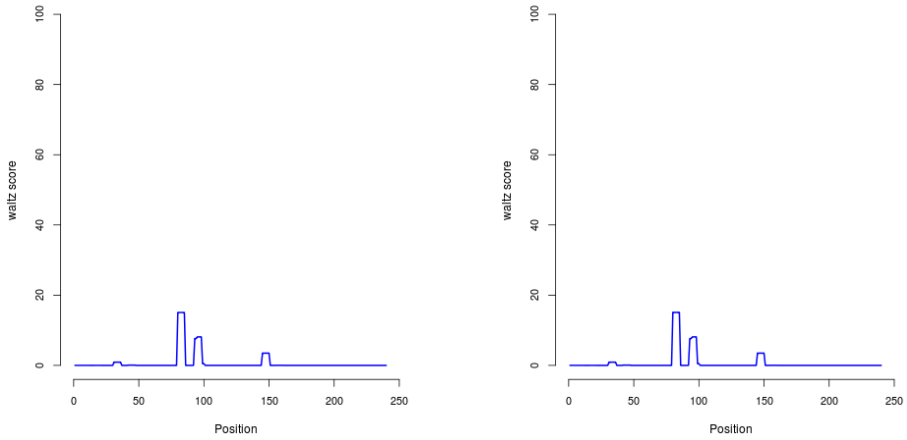

Figure 2.5: Profile representation of the WALTZ stretches in the wild type (left) and mutant (right) protein. This graph plots the per-residue WALTZ aggregation score of the wild type and variant protein. From left to right, all residue scores from the N-terminus to the C-terminus are plotted.

protein are visualized, represented by respectively a bar or profile representation. In table 2.2 , the short stretches are listed for both wild type and mutant. To compare the effect of the mutation to the WT, we also show a Difference profile (Figure 2.6), that plots the difference between WT protein and the variant.

Table 2.2: WALTZ regions in variant and wild type. For each WALTZ region, the start, end, sequence and score is given.

Number	Start	End	Stretch	Score
Wild Type				
1	79	85	GVVLFR	15.06
2	92	98	LDAFFA	8.03
Mutant				
1	79	85	GVVLFR	15.09
2	92	98	LDAFFA	8.03

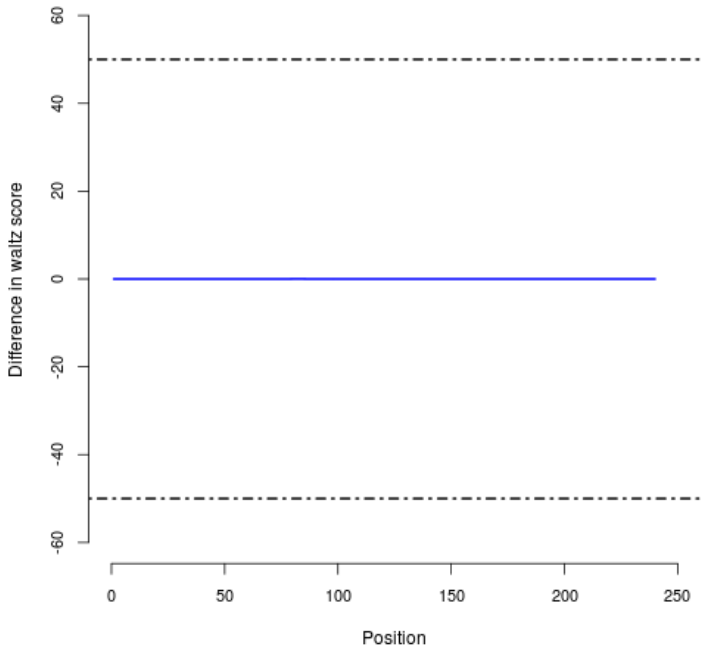

Figure 2.6: Difference in WALTZ amyloid propensity between WT and variant. This graph plots the per-residue WALTZ aggregation score diference between WT protein and the variant. From left to right, all WALTZ score differences from the N-terminus to the C-terminus are plotted. A flat line indicates that the variant does not alter the aggregation profile of the protein. Positive peaks indicate increased amyloid propensity due to this mutation. Negative peaks indicate decreased amyloid propensity due to this mutation.

2.3 Limbo prediction

LIMBO is a chaperone binding site predictor for the Hsp70 chaperones, trained from peptide binding data and structural modeling. The total LIMBO score for your protein is 2420.15 and a mutation can increase (dLIMBO >50), decrease (dLIMBO <-50) or not affect chaperone binding (dLIMBO between -50 and 50). In this case, dLIMBO equals 0.00 which means that the mutation does not affect the chaperone binding tendency of your protein.
In figure 2.7 and 2.8 the position of the LIMBO stretches in the wild type and variant protein are visualized, represented by respectively a bar or profile representation. In table 2.2, the short stretches are listed for both wild type and mutant. To compare the effect of the mutation to the WT, we also show a Difference profile (Figure 2.9), that plots the difference between WT protein and the variant.

Figure 2.7: Bar representation of the LIMBO windows present in the wild type (top) and mutant protein (bottom). In the bar representation, the position of the aggregating stretches is visualised in pink, and the dashed vertical line in the variant indicates the position of the variant residue.

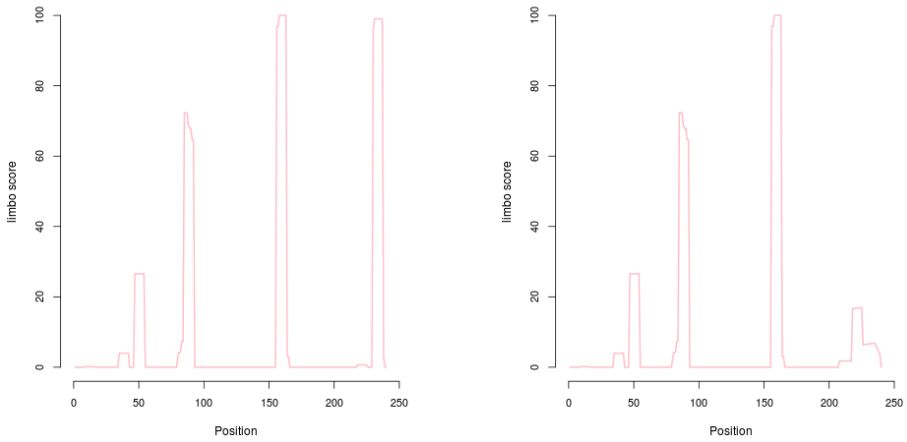

Figure 2.8: Profile representation of the LIMBO stretches in the wild type (left) and mutant (right) protein. This graph plots the per-residue LIMBO aggregation score of the wild type and variant protein. From left to right, all residue scores from the N-terminus to the C-terminus are plotted.

Table 2.3: LIMBO regions in variant and wild type. For each LIMBO region, the start, end, sequence and score is given.

Number	Start	End	Stretch	Score
Wild Type				
1	46	54	QEVMQRRD	26.56
2	82	92	LFRQLAPRAK	56.59
3	155	163	LWRYRAGL	99.62
4	229	237	KRRRESEC	99.00
Mutant				
1	46	54	QEVMQRRD	26.56
2	82	92	LFRQLAPRAK	56.59
3	155	163	LWRYRAGL	99.62
4	217	237	WERQAREHSERKKGRRESEC	10.81

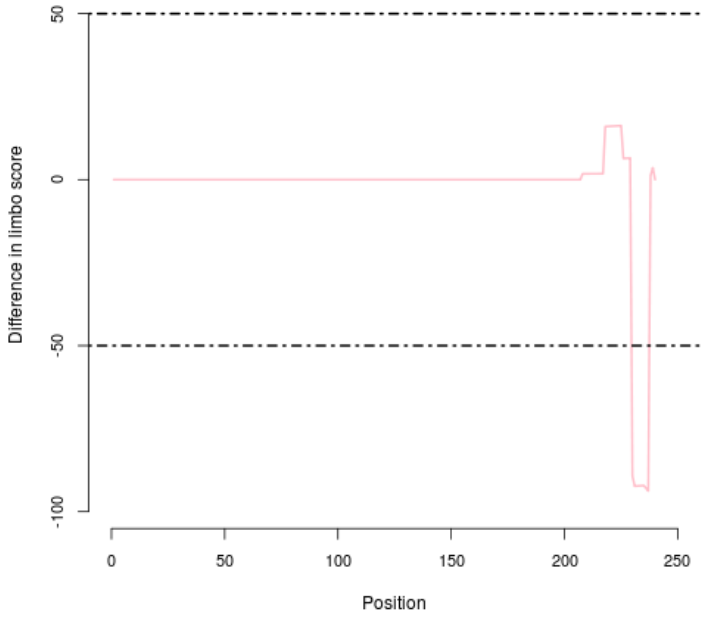

Figure 2.9: Difference in LIMBO chaperone binding propensity between WT and variant. This graph plots the per-residue LIMBO chaperone binding score diference between WT protein and the variant. From left to right, all LIMBO score differences from the N-terminus to the C-terminus are plotted. A flat line indicates that the variant does not affect the chaperone-binding sites of the protein. Positive peaks indicate increased chaperone binding due to this mutation. Negative peaks indicate decreased chaperone binding due to this mutation.

2.4 FoldX prediction

SNPeffect couldn't find reliable structural information for your protein to carry out a FoldX stability analysis. You can try to lower the homology treshold set in the submit box but keep in mind that the model accuracy tends to drop with decreasing sequence identity.

2.5 Conclusion

Finally we can conlude that:

- Based on TANGO, the mutation does not affect the aggregation tendency of your protein.
- Based on WALTZ, the mutation does not affect the amyloid propensity of your protein.
- Based on LIMBO, the mutation does not affect the chaperone binding tendency of your protein.

2.6 Explanation of all files in the .zip package

- WT_tango.png: TANGO aggregation profile score plot for the wild type sequence
- WT_tango_bar.png: Visual summary of the TANGO short stretches for the wild type sequence
- WT_waltz.png: WALTZ amylogenic profile score plot for the wild type sequence
- WT_waltz_bar.png: Visual summary of the WALTZ short stretches for the wild type sequence
- WT_limbo.png: LIMBO chaperone binding profile score plot for the wild type sequence
- WT_limbo_bar.png: Visual summary of the LIMBO short stretches for the wild type sequence
- MT_tango.png: TANGO aggregation profile score plot for the mutated sequence
- MT_tango_bar.png: Visual summary of the TANGO short stretches for the mutated sequence
- MT_waltz.png: WALTZ amylogenic profile score plot for the mutated sequence
- MT_waltz_bar.png: Visual summary of the WALTZ short stretches for the mutated sequence
- MT_limbo.png: LIMBO chaperone binding profile score plot for the mutated sequence
- MT_limbo_bar.png: Visual summary of the LIMBO short stretches for the mutated sequence
- MT_WT_tango_dif.png: Difference in TANGO aggregation between wild type and mutant
- MT_WT_waltz_dif.png: Difference in WALTZ amylogenicity between wild type and mutant
- MT_WT_limbo_dif.png: Difference in LIMBO chaperone binding between wild type and mutant