Homology modelling and comparative docking analysis of two naturally occurring pancreatic glucokinase mutants

Jose Isagani B. Janairo[1,3] and Gerardo C. Janairo[2,3]

[1]Physics Department, College of Science
[2]Chemistry Department, College of Science
[3]Center for Natural Sciences and Environmental Research (CENSER), College of Science
De La Salle University, 2401 Taft Avenue, Manila 1004, Philippines

The homology models of two naturally occurring pancreatic glucokinase mutants with contrasting enzymatic behaviours were successfully generated and accurately predicted. The homology models of the activated V367M and deactivated R369P mutants were used for comparative docking analysis in order to probe why such mutations led to either an increase or diminishment of enzyme activity. Results of structural characterization and docking simulations suggest that the small conformational changes are responsible for the observed variation in enzymatic activity. Iterative fitting and comparison of their respective Ramachandran plots reveal that these conformational changes are not sufficient to perturb the overall protein architecture. However, active site modelling showed that these conformational changes altered the manner of ligand binding, from which the observed contrasting enzymatic behaviours originate.

Email Address:
jose.isagani.janairo@dlsu.edu.ph
gerardo.janairo@dlsu.edu.ph
Submitted: November 24, 2011
Accepted: November 28, 2011
Published: January 13, 2012
Editor-in-charge: Eduardo A. Padlan
Reviewer:
Eduardo A. Padlan

KEYWORDS

homology modelling, docking calculations, pancreatic glucokinase, diabetes, single-point mutations, enzyme activity

INTRODUCTION

It has long been established that aside from environmental influences, genetic factors play a significant role in the occurrence and beginning of the pathogenesis of diabetes (Kahn et al. 1996). Among the various genetic factors that contribute to the onset of diabetes, mutations involving the pancreatic glucokinase prove to be especially problematic due to its potential association in two extreme metabolic abnormalities. The pancreatic glucokinase, dubbed as the "glucose sensor", is responsible for inducing the pancreatic β-cells to secrete insulin in response to an increase in blood glucose levels. This kinase catalyzes the phosphorylation of glucose to yield glucose-6-phosphate which is the first step and one of the rate-limiting steps in glycolysis. Mutations which lead to the inactivation of the glucokinase characterized by a decrease in V_{max} for glucose usually cause lowered insulin secretion and mild fasting hyperglycaemia (Froguel et al. 1992). Furthermore, the occurrence of mature onset diabetes of the young (MODY) which is a form of non-insulin dependent mellitus has been attributed to this type mutation. At the other extreme, abnormally high enzymatic activity of glucokinase arising from mutations oftentimes leads to congenital hyperinsulinemia (Glaser et al. 1998). Needless to say, it is of paramount importance to understand the enzymatic behaviour of mutant glucokinase at the

molecular level in order to devise specialized and personalized diabetes management programs since 620 mutations in the *GCK* gene have been reported in more than 1000 families which led to the prevalence of various glucokinase mutants (Osbak et al. 2009). Furthermore, genetic-screening studies conducted among the Italians (Massa et al. 2001), Jordanians (Khalil et al. 2009), Argentineans (Lopez et al. 2009), among others have linked the ethnic predisposition of an individual with the type of glucokinase mutation that occurs. Unfortunately, a study of similar nature has not yet been conducted in the Philippines, which will greatly contribute to the development of diabetes management protocols that especially cater to the Filipino population. Moreover, a fundamental understanding of these metabolic events can aid the development of therapeutic agents that can modulate the activity of the mutants, depending on their nature. Recently, Ding et al. (2010) conducted a univariate comparison of two naturally occurring glucokinase mutants against the wild type, namely V367M and R369P. It was observed that the former mutant variant is slightly more active with that of the wild type, in contrast to the latter mutant which exhibited a 30-fold reduction in activity.

In order to provide perspectives on these observations from a substrate-host vantage point, protein structure-based methodologies should be conducted in order to obtain an idea on the nature of interactions existing between the ligand and its corresponding receptor. However, crystal structures of the mutant variants are not available from online databases. Thus, the study presented herein provides the first account of constructing homology models for the aforementioned glucokinase mutants. The constructed homology models are the only means by which structure-based studies can be conducted in the absence of definitive crystal structures since these models are the next best thing to x-ray derived protein structures. These homology models will serve as useful tools in clarifying the observed kinetic and catalytic behaviour of the mutant variants through comparative docking calculations, thus offering beneficial contribution to the deeper understanding of the possible reasons for either the enhanced or the reduced activity of these glucokinase mutants and their probable consequences with respect to the pathogenesis of diabetes.

MATERIALS AND METHODS

Generation of Homology Models of Mutant Variants

The crystal structure and complete amino acid sequence of the human wild-type pancreatic glucokinase was obtained from the Protein Data Bank (PDB) through its PDB ID which is 3IDH (Petit et al. 2011). The wild-type structure was used as the template for the generation of homology models utilizing the software Modeller (Sali and Blundell 1993) version 9.9. Details of the standard procedure for homology model generation, such as required run files, command syntaxes and expected outputs, can be found at the software manual. Five models for each mutant variant were generated, from which the model with the lowest discrete optimized protein energy (DOPE) score was chosen as the best model and used for further analysis.

Comparative Structural Characterization of Mutant Variants with the Wild-Type Glucokinase

The structural features of the best homology models were then compared with those of the wild-type using the various functions of the Deep View / Swiss-Pdb Viewer version 4.0.1 program (Guex and Peitsch 1997). Comparison of the secondary and tertiary structures was carried out by superimposing the wild-type against each of the mutant variants and thereafter calculating their respective relative mean standard deviation. Fitting was based on the alignment of the α-carbons as well as that of the backbones of the proteins. Furthermore, their corresponding Ramachandran plots were calculated in order to obtain a holistic structural comparison and characterization of the wild-type and mutant enzymes.

Substrate Geometry Optimization

The substrate used for the molecular docking calculations was α-D-glucopyranose. The reason for the choice of the substrate is to maintain consistency since this type of glucose is co-crystallized with that of the wild type pancreatic glucokinase used in this study. The α-D-glucopyranose was drawn using the molecular builder function of ArgusLab 4.0.1 (Thompson 2004) and its initial geometry was refined using molecular mechanics calculations through the universal force field (UFF). From the refined structure, geometry optimization was then conducted through a restricted shell semi-empirical calculation using the AM1 Hamiltonian. Approximately 500 cycles were needed to attain convergence for the energy minima.

Comparative Molecular Docking

The wild-type PDB model and the best homology models for the two mutant variants were used as the target proteins for the molecular docking calculations wherein the previously geometry-optimized glucose served as the ligand. The Argus Dock function of Argus Lab was utilized for the docking simulations. The glutamate residue at the 256 position was selected as the binding site, restricting conformational search within a 15x15x15 Angstrom volume centred on this residue. The Glu256 residue is responsible for binding the glucose ligand (Gidh-Jain et al. 1993). Grid resolution was set at 0.4 and torsional flexibility was enabled in order to take into account the various degrees of freedom that the ligand and receptor may adopt during docking. The pose with the lowest energy was chosen for visual analysis of ligand-receptor interactions.

RESULTS AND DISCUSSION

Structure-based modelling of proteins requires a reliable crystal structure since it is where the majority of the architectural information is contained, such as ligand orientation, conformation in solution, side-chain and backbone angles, etc. In the absence thereof, or of 2D NMR coordinates, homology models are adequate substitute, if they had been constructed in a satisfactory manner. A homology model is a predicted, or

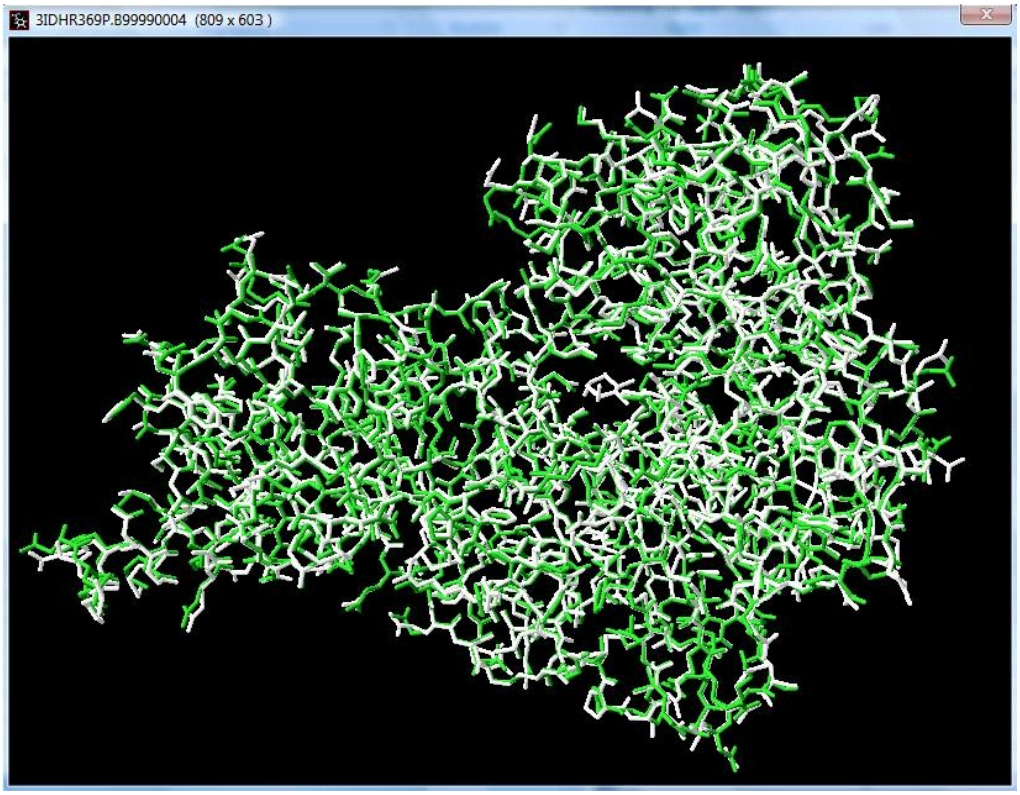


Figure 1. Superposition of R369P mutant (green) with WT (white).

approximated, 3D structure of a protein based on its alignment with a template (Young 2009). The choice of which template to use usually depends on the amino acid sequence and the predominant secondary and tertiary structures. If the protein to be modelled belongs to a family wherein other family members have already been crystallized, the template to be used is the family member exhibiting the highest degree of amino acid sequence similarity. For mutants on the other hand, the most logical choice of template is the wild-type protein since mutants usually differ by only a few amino acid residues. In our study, the template used was the wild-type pancreatic glucokinase. The number of homology models to be yielded after the alignment was set to five, wherein the best homolog is judged as the model exhibiting the lowest DOPE scores and GA341 value closest to 1.

The DOPE value is a statistical function dependent on the distances of the atoms of the target protein to those of the template. The DOPE score assigns values based on the position and orientation of all non-hydrogen atoms wherein a low number represents an accurate model prediction. The GA341 parameter, on the other hand, assesses the folding state of the generated model. Possible values for this criterion is from 0 to 1 wherein a poorly predicted fold-model is given a 0.0 score and a well-approximated model yields a 1.0 (Eswar et al. 2008). In our generated models, model 4 and model 5 were judged to be the best models for the R369P and V367M mutants, respectively (Table 1). All generated models exhibited a GA341 score of 1.0, which is an indication of the reliability of the generated models.

In order to determine if the single point substitution for both mutant variants resulted in a significant perturbation of the secondary and tertiary structures from those of the wild type, iterative alpha-carbon and backbone fitting, as well as comparison of their corresponding Ramachandran plots, were conducted.

Results of iterative fitting reveal that the mutants adopt a very similar overall structure with that of the wild-type as seen from both the visual superposition (Figures 1 and 2) and the calculated relative mean standard deviation (Table 2). In support of the previous findings, a quick inspection of the Ramachandran plots of the three enzymes (Figure 3) indicates likewise. Taken the structural data as a whole, the small RMSD values (<1) and nearly identical Ramachandran plots suggest that any conformational change the mutation might induce is localized only within the close proximity of the involved residues. These conformational changes that arise from the accompanied side-chain variation are not enough to initiate significant structural perturbation for both mutants.

The various interactions that arise between the ligand and the enzymes must be taken into account in order to satisfactorily explain the observed activating and deactivating nature of the glucokinase mutants. The accurately predicted homology models for both mutants will also serve as the target proteins for the comparative docking analysis in order to quantitatively and qualitatively analyze the interactions existing between glucose and the enzymes. The geometry of the glucose ligand was optimized using semi-empirical quantum chemical calculations in order to obtain reliable geometrical information with respect to the structure of the ligand. By ensuring that the structure of the ligand is properly optimized, its interaction with the environment of the active site can be thus sufficiently scrutinized. Among the numerous docking softwares that are available either commercially or free by virtue of academic licensing, Argus Lab was chosen due to the reliability of its calculation for binding free energies (Oda and Takahashi 2009) as well as pose construction (Oda et al. 2007). Results of the comparative docking simulations depict that the glucose ligand

adopts a very different spatial pose within the active sites of the two mutants relative to the wild-type glucokinase.

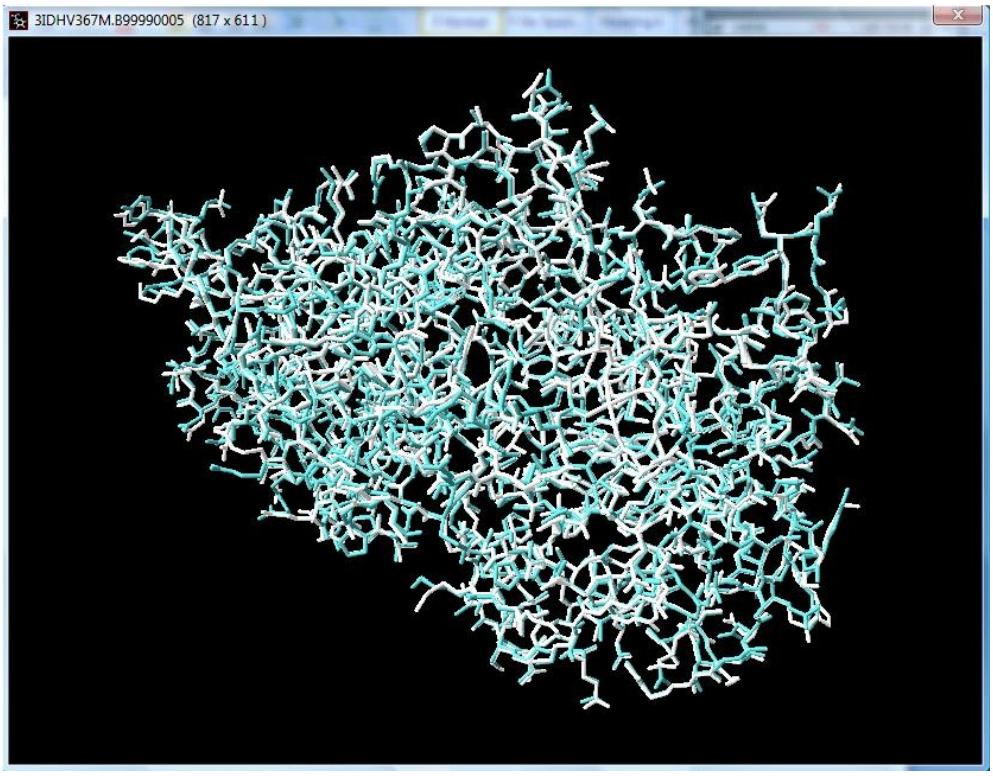


Figure 2. Superposition of V367M mutant (cyan) with WT (white).

Aside from spatial orientation, another striking difference between the wild-type and the two mutants is the number of H-bonds that were formed between the active site residues and the ligand. It is highly possible that such point of difference can explain the increased activity of the V367M mutant and the deactivated nature of the R369P mutant variant. It is a well established principle that multiple non-covalent and non-repulsive interactions between the natural ligand and its receptor lead to transition state stabilization which ultimately results in rate enhancement (Shan and Herschlag 1996). However, this principle is not absolute and it is plausible to assume that such is not always the case when the ligand is not the natural substrate of the enzyme and/or the protein of interest is a mutant variant. In the case of the first scenario, the model for competitive inhibition can be applied wherein an inhibitor that closely resembles the property and structure of the natural substrate competes for the binding of the active site. However, multiple interactions emanating from the active site and the inhibitor do not lead to rate enhancement but rather enzymatic incapacitation. As a matter of fact, the more non-covalent interactions there are the better inhibition results. For the second case which is more applicable to the presented model, mutations in whatever form and diversity can lead to slight conformational changes within the protein. The changes may not be enough to modify the overall structure of the protein but it may be sufficient to alter the behaviour of the enzyme, as what has been previously discussed. For instance, a conservative mutation from isoleucine to leucine drastically changed the metal cofactor preference of *EcoRV* from magnesium to manganese (Vipond et al. 1996). The effect of these small conformational changes can be seen in the case of the two mutants wherein different residues in the active site participate in interacting with the glucose ligand compared to the wild type (Table 3). To be more specific, Glu290 is involved in V367M but not with the wild-type; similarly for Gln286 of R369P. Thus, these conformational changes are responsible for the observed catalytic modification. As in the case of the deactivated R369P mutant, it is possible that these conformational changes give rise to a competition between thermodynamic and kinetic control wherein the former predominates over the latter. The most stable ligand orientation, although more thermodynamically favourable may orient the ligand in such a way that is not optimum for the reaction. The multiple interactions which stabilize the system may "tie" the ligand in an orientation wherein its reactive centers may be far away from the catalytic residues of the active site. Although more stable, the ligand may need to pay torsional penalties just to re-align itself in such a way that is optimum for the reaction. This torsional penalty may offset whatever stabilization the multiple interactions may provide, as well as introduce a time delay due to its re-positioning. These factors may explain the observed decrease in activity for the R369P mutant variant. For the activated mutant on the other hand, the converse of the previous situation may hold true. It is possible that the ligand, by virtue of the conformational changes that have been established by the mutation, has been positioned in such a way that better suits the reaction. In this manner, both thermodynamic stabilization and kinetic control contribute to the enhancement of the activity of the mutant.

CONCLUSION

In summary, the homology models of two naturally occurring mutants of pancreatic glucokinase have been accurately generated. The homology models were then used as target proteins for the comparative docking analysis in order to explain the enhanced activity of V367M and the incapacitated R369P mutants. Simulations suggested that the small

conformational changes that arose from the mutations led to the two dissimilar observations. In the case of R369P, the conformational changes led to multiple interactions that "tied" the glucose ligand in such a way that is not favourable to reaction, in contrast to V367M wherein its enhanced activity may be derived from its proper alignment and orientation within the active site which was led by the aforementioned conformational changes. The findings presented are useful in the design of either activators and inhibitors of these types of glucokinase mutants as well as in the understanding of the interactions that are in play in these specific model systems.

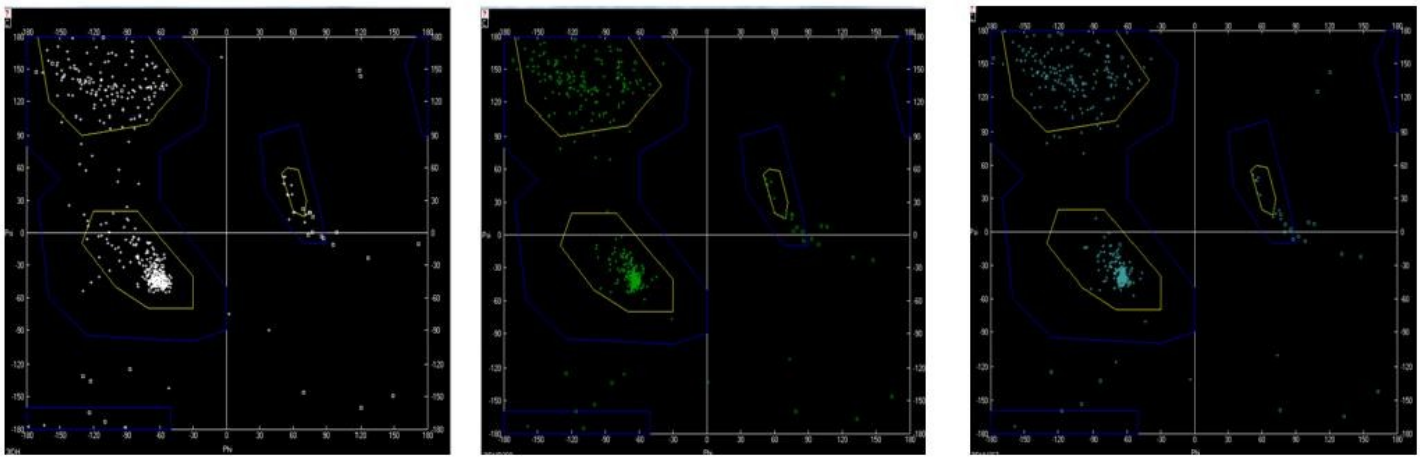
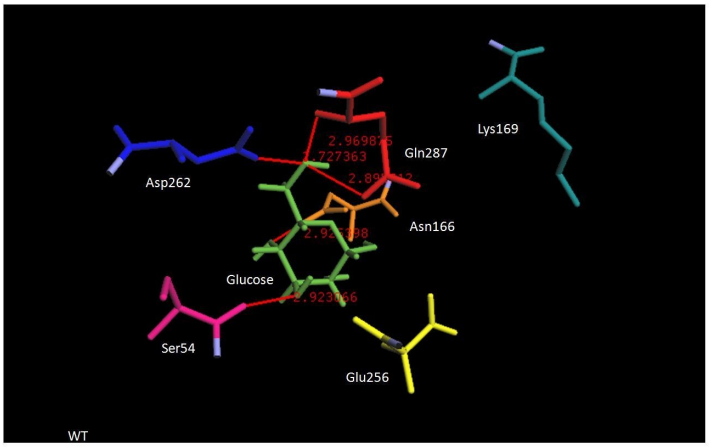
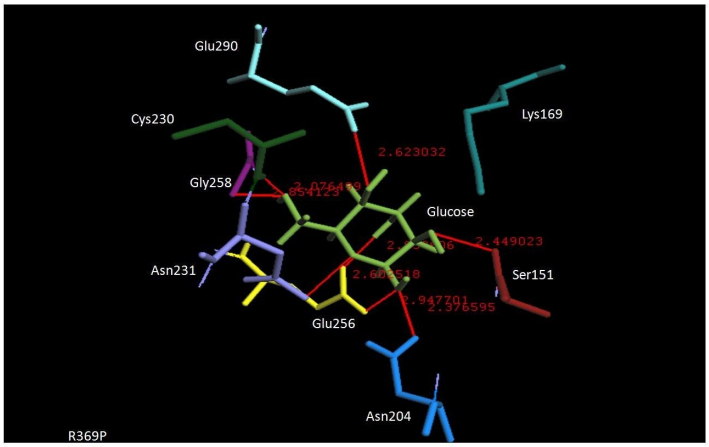


Figure 3. Comparative analysis of the Ramachandran plots of the WT (white), R369P (green) and V367M (cyan) mutants.





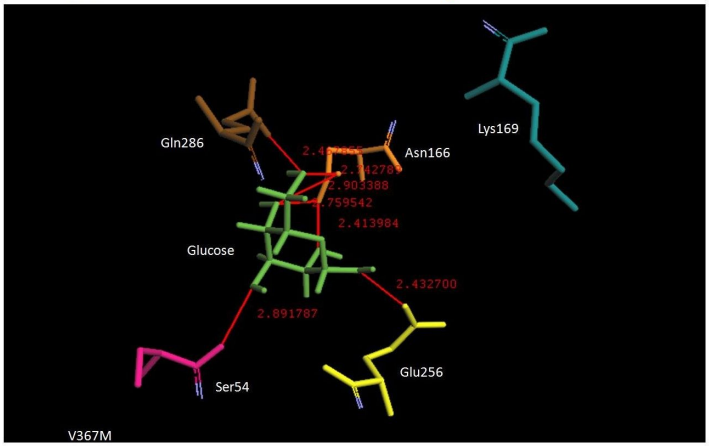


Figure 4. Lowest energy pose for glucose with WT (top).

Figure 5. Lowest energy pose for glucose with R369P mutant (upper right).

Figure 6. Lowest energy pose for glucose with V367M mutant (lower right).

CONFLICT OF INTEREST

The authors certify and declare there are no conflicts of interest associated with the presented study.

Table 1. Summary of DOPE model scores for the two glucokinase mutants. Highlighted are the scores of the two models used in further analyses.

Model	DOPE Model Scores	
	R369P	V367M
1	-55891.09375	-55698.20703
2	-55834.95703	-55771.63672
3	-55271.46484	-55668.70703
4	-56076.01172	-56075.88672
5	-55833.08984	-56216.65625

Table 2. Relative mean standard deviation (RMSD) after fitting the mutant variants against the wild-type with respect to the alpha-carbon and the protein backbone

Point of fitting	RMSD	
	R369P	V367M
Alpha-carbon	0.18	0.14
Protein backbone	0.24	0.20

Table 3. Summary of quantitative data of comparative docking analysis

Enzyme	Lowest Energy Pose (Kcal / mol)	Highest Energy Pose (Kcal / mol)	Number of Plausible Poses Found	Residues Involved	H-bonds
WT	-6.81	-1.65	76	4	5
V367M	-7.3	-1.21	83	4	7
R369P	-6.97	-0.38	150	7	8

REFERENCES:

Ding S-Y, Tribble ND, Kraft CA, Markwardt M, Gloyn AL, Rizzo MA. Naturally occurring glucokinase mutations are associated with defects in posttranslational S-nitrosylation. Mol Endocrinol 2010; 24: 171-177.

Eswar N, Eramian D, Webb B, Shen M-Y, Sali A. Protein structure modeling with MODELLER. Methods Mol Biol 2008; 426: 145-159.

Froguel PH, Vaxillaire M, Sun F, Velho G., Zouali H, Butel MO, Lesage S, Vionnet N, Clement K, Fougerousse F, Tanizawa Y, Weissenbach J, Beckamnn JS, Lathrop G.M, Pasa PH, Permutt MA, Cohen D. Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus. Nature 1992; 356: 162-164.

Gidh-Jain M, Takeda J, Xu LZ, Lange AJ, Vionnet N, Stoffel M, Froguel P, Velho G., Sun F, Cohen D, Patel P, Lo Y-MD, Hattersley AT, Luthman H, Weddel A, St. Charles R, Harrsion RW, Webber IT, Bell GI, Pilkis SJ. Glucokinase mutations associated with non-insulin-dependent (type 2) diabetes mellitus have decreased enzymatic activity: implications for structure / function relationships. Proc Natl Acad Sci USA 1993; 90: 1932-1936.

Glaser B, Kesavan P, Heyman M, Davis E, Cuesta A, Buchs A, Stanley CA, Thornton PS, Permutt MA, Matschinsky FM, Herold KC. Familial hyperinsulinism caused by an activating glucokinase mutation. N Engl J Med 1998; 338: 226-230.

Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 1997; 18: 2714-2723.

Kahn CR, Vicent D, Doria A. Genetics of non-insulin-dependent (Type II) diabetes mellitus. Annu Rev Med 1996; 47: 509-531.

Khalil R, Al-Sheyab F, Khamaiseh E, Halaweh MA, Abder-Rahman HA. Screening of mutations in the *GCK* gene in Jordanian maturity-onset diabetes of the young type 2 (MODY2) patients. Genet Mol Res 2009; 8: 500-506.

Lopez AP, Foscaldi SA, Perez MS, Krochik G., Rodriguez M, Traversa M, Puchulu FM, Hirschler V, Bergada I, Frechtel GD. Glucokinase gene mutation screening in Argentinean clinically characterized MODY patients. Exp Clin Endocrinol Diabetes 2009; 117: 391-394.

Massa O, Meschi F, Cuesta-Munoz A, Caumo A, Cerutti F, Toni S, Cherubini V, Guazarrotti L, Sulli N, Matschinsky M, Lorini R, Iafusco D, Barbetti F. High prevalence of glucokinase mutations in Italian children with MODY. Influence on glucose tolerance, first-phase insulin response, insulin sensitivity and BMI. Diabetes Study Group of the Italian Society of Paediatric Endocrinology and Diabetes (SIEDP). Diabetologica 2001; 44: 898-905.

Osbak KK, Colclough K, Saint-Martin C, Beer NL, Bellanne-Chantelot C, Ellard S, Gloyn AL. Update on mutations in glucokinase (*GCK*), which cause maturity-onset diabetes of the young, permanent neonatal diabetes, and hyperinsulinemic hypoglycaemia. Hum Mutat 2009; 30: 1512-1526.

Oda A, Takahashi O. Validation of ArgusLab efficiencies for binding free energy calculations. Chem-Bio Inf J 2009; 9: 52-61.

Oda A, Okayasu M, Kamiyama Y, Yoshida T, Takahashi O, Matsuzaki H. Evaluation of docking accuracy and investigation of roles of parameters and each term in scoring functions for protein-ligand docking using ArgusLab software. Bull Chem Soc Jpn 2007; 80: 1920-1925.

Petit P, Antoine M, Ferry G, Boutin JA, Lagarde A, Gluais L, Vincentelli R, Vuillard L. The active conformation of human glucokinase is not altered by allosteric activators. *Acta Crystallogr, Sect D* 2011; 67(Pt 11): 929-935.

Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993; 234: 779-815.

Shan S, Herschlag D. Energetic effects of multiple hydrogen bonds. Implications for enzymatic catalysis. J Am Chem Soc 1996; 118: 5515-5518.

Thompson, MA. ArgusLab 4.0.1. Planaria Software LLC, Seattle, WA, 2004.

Vipond IB, Moon B-J, Halford SF. An isoleucine to leucine mutation that switches the cofactor requirement of the *Eco*RV restriction endonuclease from magnesium to manganese. Biochemistry 1996; 35: 1712-1721.

Young DC. Computational drug design: A guide for computational and medicinal chemists. New Jersey: John Wiley & Sons, 2009: 43.