Investigating Mutations in G6PD

Leslie MacMillan*[a]

Received (2/5/2014)

Abstract

G6PD is an enzyme important to the blood system. When the enzyme is deficient, red blood cells lyse in the presence of oxidative damage. This can manifest itself in a variety of ways, from anemia to jaundice[9]. G6PD deficiencies are the most common enzyme deficiency in the world [12]. These deficiencies may have been selected for due to ability to counteract malaria. The deficiency of the enzyme can be caused by many different mechanisms. The degree of severity of the phenotype of the disorder is usually dependent on the degree of functional change in the protein. In this paper we explore the various classes of mutations, and the way they affect the structure of the protein and ultimately its function.

Introduction

G6PD is a widely studied protein that has many mutations. Its OMIM entry lists 63 examples. Serious problems with this gene can be fatal. There is a range of effects resulting from various levels of variance from the true protein. In this paper, we will attempt to classify these mutations by their effects on the protein's structure and function.

G6PD deficiency has been shown to be selected for in certain areas of the world, likely due to its reduction of susceptibility to malaria[9]. It has even been used to study migration patterns over human history due to the way it may have evolved independently among certain groups[3], and the genetic traces these evolutions leave.

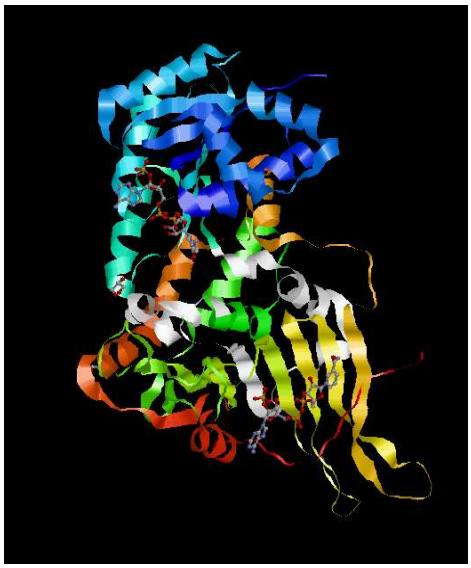

Figure 1. PDB File 2BH9 with active site obtained from Catalytic Site Atlas in White.

Results and Discussion

Identification of Active Site

The active site appears to be between residue Asp200 and His 263[6]. The hypothesis is that a mutation in this area would lead to functional problems with the protein, and the extent of the functional problem would depend on the extent of conformational change in the active site caused by the mutation.

In figure 1, the active site is shown in white. The non-protein molecules in the figure are NADP+ molecules.

SAAPdap Analysis

Using the resource, Single Amino Acid Polymorphism data analysis pipeline[7], those mutations which only involved one amino acid were entered one at a time into the analyser. Sample results are seen in the table. The first two entries are shown as they are located in the physical structure of the non-diseased variant in figures 2 and 3.

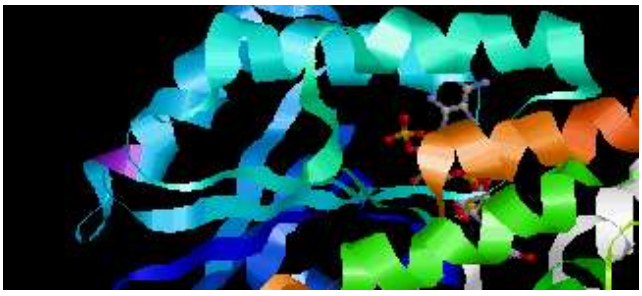

Figure 2. Residue 126 is coloured purple. It has no structural effects according to SAAPdap, and does not appear to be close to the NADP+ binding site or the active site (white). However, in vivo, the mutation causes the protein to be unstable.

Mutant	Structures in db	Structural Effects	Phenotype Name	Disease Effects and Notes
ASN126ASP	11	No Structural Effects Identified	G6PD A+	Unstable protein in vivo
ALA335THR	11	CorePhilic (2) Impact (11)	G6PD CHATHAM	Class 2 Enzyme derangement
GLU156LYS	11	No Structural Effects Identified	G6PD ILESHA	Class 3 mutation
GLY163SER	11	Clash (3) Interface (2)	G6PD MAHIDOL	In loop region. Causes class 2 enzyme derangement Reduces p. vivax parasite density
SER188PHE	11	Surface Phobic (11)	MEDITERANNEAN	Causes class II abnormality
ASP282HIS	11	No structural effects	SEATTLE-LIKE	Mild Effects
ARG393HIS	11	Binding (6) Interface (9)	PORTICI	Native residue involved in interface/binding Causes nonspherocytic hemolytic anemia
ASN363LYS	11	HBonds(11)		Causes nonspherocytic hemolytic anemia
GLU317LYS	11	No structural effects identified	KERALA-KALYAN	Changes charge of the molecule. Mild effects
ARG198CYS	11	BuriedCharge(11) Hbonds(3) Impact(11)	COIMBRA	Causes deficiency

Table 1: Information from OMIM and SAAPdap.

SAAP Analysis

UniProt Entry: P11413
Mutation: ARG 198->CIS

Summary

PDB	chain	HBonds	Binding	BuriedCharge	Voids	SProtFT	SurfacePhobic	Interface	Clash	CisPro	CorePhilic	Impact	SSGeom
1gki	A			X								X	
1gki	B			X								X	
1gki	C			X								X	
1gki	D			X								X	
1gki	E			X								X	
1gki	F			X								X	
1gki	G			X								X	
1gki	H			X								X	
2bh9	A	X		X								X	
2bhl	A	X		X								X	
2bhl	B	X		X								X	

Figure 2. Sample SAAP analysis table of effects. Impact indicates the change was at a highly conserved site..

Figure 3. Residue 335 is the small red colored section in the green area. SAAPdap raised warnings that it is in a highly conserved region and makes a hydrophobic region hydrophilic. This may result in structural changes in the protein. It can be seen that this is close to the NADP binding site, so it could have functional effects.

Sample SAAPdap results are given in Table 1, annotated with information from the OMIM entries.

Drastic effects such as non-spherocytic anemia, were observed in mutations that occurred at the binding site, and in one that disrupted a hydrogen bond that ocurred at the intersection of two domains.

Mild effects were often associated with no structural changes in the molecule. However, there were often at least mild effects, presumably from an influence smaller than, for-example, non-binding of the ligand. For example, one mutation was associated with a change in charge of the molecule, reducing the enzyme's efficacy, but not as significantly as other mutations did.

CATH analysis

When superimposed, the general shape of the proteins in the family seemed to have a similar shape around the middle, with variation in outer regions, indicating the active site was somewhere in the center of the molecule[10], which was confirmed by the catalytic site atlas.

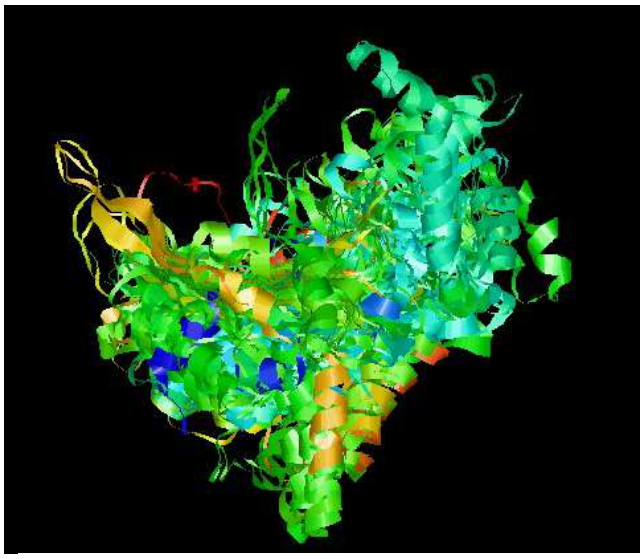

Figure 4: Superposition of structures in superfamily 3.30.360.10, Dihydrodipicolinate Reductase, gives impression of conserved structures.

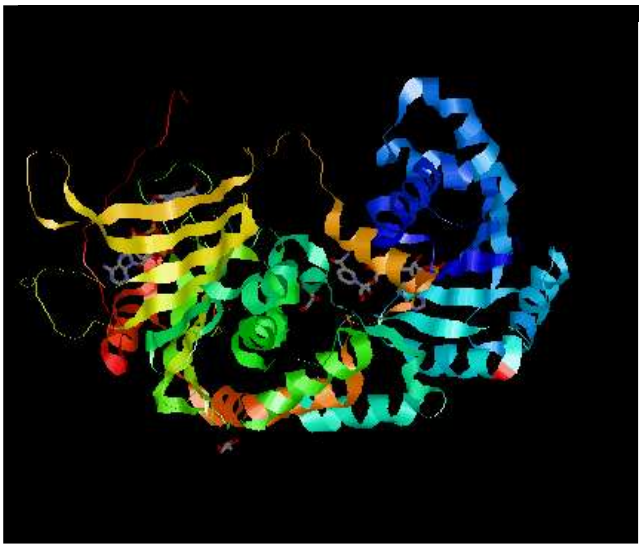

Figure 5. Compare superposition to G6PD structure

Materials and Methods

First, some initial research into the effects of G6PD and its deficiency was performed. Then the OMIM listing for protein was investigated. Numbering uniformity was confirmed between UniProKB sequence and OMIM sequence (found to be the same).

The catalytic site atlas was searched for the active site, found to be located on residues 200-263[6].

Then a three-dimensional structure from PDB was downloaded. At first glance, it seemed that PDB had only obtained structures for variants of G6PD, and it seemed it was not possible to see the active site of normal G6PD interacting with its ligand. Therefore, the paper providing this structure was examined. Upon further investigation it was seen that the variant was designed for reasons related to the crystallization of the protein[8]. In fact, the residues that were deleted were poorly conserved and therefore unlikely to be functionally significant[8]. So this showed that PDB contained a relevant 3D structure which was then observed and its display manipulated, in RasMol.

SAAPdap

Potential structural effects of Single Amino Acid Polymorphisms were predicted using the SAAPdap tool[7]. Results are summarized in table 1.

CATH

CATH provided information about the superfamily to which the protein belonged. A superposition of the structures of this superfamily is shown in figure 4.

Gene3d

Gene3d database was consulted, however did not add relevant information beyond what was supplied by CATH.

Conclusions

While mutations in or near the binding site caused G6PD deficiencies, and often the most severe kinds, it was not only these mutations which had an effect on the phenotype. Mutations in residues involved in hydrogen bonds had a drastic effect, and some that caused no structural change other than a change in charge of the molecule also had an effect. This shows that mutations can affect the function of the protein in many ways. These ways could be called unpredictable, but with the advent of computational biology and careful databank record keeping and annotation, hopefully the scientific community will be increasingly able to predict these ways in the future.

Notes and references

[a] *UCL London, UK. E-mail: leslie.macmillan.13@ucl.ac.uk*

1. Frank JE. Am Fam Physician. 2005 Oct 1;72(7):1277-82.

Diagnosis and management of G6PD deficiency.

2. Glucose-6-phosphate dehydrogenase deficiency. (2014, April 26). In *Wikipedia, The Free Encyclopedia*. Retrieved 12:06, May 1, 2014, fromhttp://en.wikipedia.org/w/index.php?title=Glucose-6-phosphate_dehydrogenase_deficiency&oldid=605844152

3. OMIM
Online Mendelian Inheritance in Man, OMIM®. Johns Hopkins University, Baltimore, MD. MIM Number: {MIM number}: {Date last edited}: . World Wide Web URL:http://www.omim.org/entry/305900?search=Glucose-6-phosphate%20dehydrogenase&highlight=glucose%20phosphate%20glucose6phosphate%20dehydrogenase

4. UniProtKB http://www.uniprot.org/uniprot/P11413

5.http://www.uniprot.org/blast/?about=P11413[2-515]

6. Catalytic Site Atlas https://www.ebi.ac.uk/thornton-srv/databases/CSA/SearchResults_UNI.php?UNIID=P11413&SUBMIT_UNI=SEARCH+CSA

7. SAAPdap, Single Amino Acid Polymorphism data analysis pipeline
Nouf S Al-Numair andAndrew CR Martin. The SAAP pipeline and database: tools to analyze the impact and predict the pathogenicity of mutations, BMC Genomics. 2013; 14(Suppl 3): S4. Published online May 28, 2013. doi:10.1186/1471-2164-14-S3-S4
URL: http://www.bioinf.org.uk/saap/dap/

8. Structural studies of glucose-6-phosphate and NADP+ binding to human glucose-6-phosphate dehydrogenase.
Kotaka, M.,Gover, S.,Vandeputte-Rutten, L.,Au, S.W.N., Lam, V.M.S.,Adams, M.J.
Journal:(2005) Acta Crystallogr.,Sect.D 61: 495
PubMed:15858258

9. Beutler, Ernest Glucose-6-phosphate dehydrogenase deficiency: a historical perspective
doi: 10.1182/blood-2007-04-077412Blood January 1, 2008 vol. 111no. 1 16-24

10. New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures.
Sillitoe I, Cuff AL, Dessailly BH, Dawson NL, Furnham N, Lee D, Lees JG, Lewis TE, Studer RA, Rentzsch R, Yeats C, Thornton JM, Orengo CA
Nucleic Acids Res. 2013 Jan
Pubmed: 23203873

11. Nucleic Acids Res. Jan 2012; 40(D1): D776–D782.
Published online Oct 17, 2011. doi: 10.1093/nar/gkr852