Chapter – 2

REVIEW OF LITERATURE

Chapter – 2

REVIEW OF LITERATURE

2.1. Glucose-6-Phosphate Dehydrogenase enzyme

The enzyme G6PD (monomer) is composed of two major domains, the smaller of which is the N-terminal domain formed by amino acids 27-200, and forms an α/β domain (Rossman fold). The co-factor ($NADP^+$) which is used in catalysis binds to this domain. The amino acids 201-515 form the C-terminal domain (larger domain). The C-terminal residues consist of nine anti-parallel strands of β-sheets, forming the α+β domain. An α-helix links the two domains, in which lies the conserved eight-residue sequence (198-206), required for the binding of substrate, Glucose-6-Phosphate. There is also an additional structural $NADP^+$ binding site, necessary for structural stability of G6PD (Mason *et al*., 2007). It has been established that an active G6PD exists in either the dimeric or tetrameric form (Figure 4).

2.2. Glucose-6-Phosphate Dehydrogenase deficiency

2.2.1. Global prevalence of Glucose-6-Phosphate Dehydrogenase deficiency

A World Health Organization (WHO) report states that 2.9% of the world's population is G6PD deficient and 7.5% of the population is carriers of the condition. From geographic point of view, places where populations have historically been exposed to endemic malaria are correlated with the prevalence of G6PD deficiency (Nkhoma *et al*., 2009). The Sub-Saharan Africa holds the record of highest prevalence of G6PD deficiency, followed by Arabian subcontinent, Central and Southeast Asia, Mediterranean, Europe and Latin American countries (Nkhoma *et al*., 2009; Howes *et al*., 2012).

2.2.2. Glucose-6-Phosphate Dehydrogenase deficiency in India

The first incidence of G6PD deficiency in India was reported in 1963 (Baxi *et al*., 1963). According to the World Health Organization, the prevalence of G6PD deficiency in India ranges from 0–10% depending on the population (WHO, 1989). Various studies on the prevalence of G6PD deficiency have reported it to range from 0% to 27% among various caste, ethnic and linguistic groups of India (Tripathy and Reddy, 2007). Studies also revealed higher frequency of G6PD deficiency among the tribal/scheduled caste population (Tripathy & Reddy, 2007; Rai & Kumar, 2012). Although detailed study has not been performed in many parts of the country but its prevalence is higher in Northern and Western parts of India as compared to other parts of the country (Bhasin, 2006). In the western part of India, the Vataliya Prajapati community has the highest prevalence of the deficiency, followed by Gonds, Warli, Parsis, Madia and Kutchi Bhanushalis (Kumar *et al*., 2016). In the Northern part, Meghwal-Chamars and Punjabi Khatris have higher frequencies of the deficiency. The Angami Nagas has the highest prevalence of G6PD deficiency in the eastern part of the country, followed by the Adi, Apatani, Nishi, Rabha, Mikir (Karbi), Santhal. The tribals of Andhra Pradesh (13%) and Kurumba tribe of Kerala (11.9%) have the highest prevalence in Southern part of India (Balgir, 2006). Till date, the highest prevalence (30.4%) of G6PD deficiency in India is reported from the Dhelki Kharia subtribe in Orissa (Kumar *et al*., 2016). Over 400 G6PD variants have been identified and these variants are specific to different geographic locations (Nkhoma *et al*., 2009).

2.3. Glucose-6-Phosphate Dehydrogenase variants

2.3.1. Glucose-6-Phosphate Dehydrogenase variants identified till date

More than 400 variants of G6PD have been reported worldwide; however, only 230 mutations have been characterized at the molecular level (Luzzatto *et al*., 2020). Based on the clinical and biochemical phenotypes these mutations are categorized into five classes, class I to V (WHO, 1989) (Table 1). A number of G6PD variants have been reported that are specific to different geographic locations. There are several mutations reaching gene frequency between 1-70% in certain populations which are regarded as polymorphic variants. G6PD variants like Gaohe, Honiara, Aures, Mahidol, Coimbra, Seattle,

Montalbano, Viangchan, Cassano, Union, Canton, Kaiping are some of the major polymorphic variants (Mason *et al*., 2007). Among all the identified variants, the Mediterranean is the most widespread variant, common in the Mediterranean areas (Spain, Italy and Greece), the Middle East and the Indian subcontinent (Vives-Corrons *et al*., 1990). This mutation results from a point mutation (C→T) at nucleotide 563. In addition, a silent mutation at nucleotide 1311 in individuals from the Mediterranean region and the Middle East is also reported, but this additional mutation is not seen in case of Mediterranean variant from India. This indicates an independent origin of a single variant in different region or population (Kurdi-Haidar *et al*., 1990).

2.3.2. Glucose-6-Phosphate Dehydrogenase variants in Southeast Asia

In Southeast Asia, significant regional variation in the predominant variants is observed, and this variation seems to be intimately linked to geographic regions as well as distinct racial and ethnic groupings (Li *et al*., 2015). Although G6PD Viangchan and G6PD Mahidol are the two most common variants identified in Southeast Asian countries, every country has a different distribution of G6PD mutations. The mutations reported from the Southeast Asian countries are presented in Table 6.

Table 6. G6PD variants reported from Southeast Asia.

Sl. No.	**Exon No.**	**Variant**	**Mutation**	**Codon**	**Class**	**Country**	**References**
1	2	Vietnam 1	Glu→Lys	2	*NR*	Vietnam	Hue *et al*., 2009
2	2	*Name unknown*	Arg→Trp	9	*NR*	Singapore	Hamada *et al*., 2010
3	2	Gaohe	His→Arg	32	III	China, Thailand, Vietnam	Chao *et al*., 1991; Phompradit *et al*., 2011; Matsuoka *et al*., 2007
4	2	Gidra	Met→Thr	37	*NR*	China	Vulliamy *et al*., 1997

5	3	Orissa	Ala→Gly	44	III	Malaysia	Ainoon *et al*., 2003
6	3	Aures	Ile→Thr	48	III	Laos	Sanephonasa *et al*., 2021
7	4	Kozukata	Trp→Cys	53	I	China	Vulliamy *et al*., 1997
8	4	Kamogawa	Arg→Trp	57	II	China	Vulliamy *et al*., 1997
9	4	Songklana-garind	Phe →Ile	66	II	Thailand	Lasombat *et al*., 2005
10	4	Vietnam 2	Phe→Cys	66	II	Vietnam	Hue *et al*., 2009
11	4	Asahi	Val→Met	68	III	China	Hirono *et al*., 2002
12	4	Murcia	Tyr→Cys	70	III	Singapore	Hamada *et al*., 2010
13	4	Vietnam 3	Ser→Ser	73	*NR*	Vietnam	Hue *et al*., 2009
14	5	Guangzhou	Pro→Ser	92	III	China	Jiang *et al*., 2006
15	5	Bao Loc	Trp→His	118	II	Vietnam	Matsuoka *et al*., 2007
16	5	Vanua Lava	Leu→Pro	128	II	Indonesia, Thailand, Malaysia, Phillipines	Sulistyaningrum *et al*., 2020; Nuinoon *et al*., 2022; Ainoon *et al*., 2003; Minucci *et al*., 2011
17	5	Quing Yan / Chinese-4	Gly→Val	131	III	China, Thailand, Vietnam,	Chiu *et al*., 1993; Phompradit *et al*., 2011; Matsuoka

						Singapore, Laos	*et al*., 2007; Hamada *et al*., 2010; Sanephonasa *et al*., 2021
18	5	Liuzhou	Glu→Lys	148	II	China	Yan *et al*., 2006
19	5	Shenzen	Cys→Tyr	158	II	China	Chen *et al*., 2010
20	6	Mahidol	Gly→Ser	163	III	Myanmar, Thailand, Malaysia, Vietnam, Laos, Cambodia	Lee *et al*., 2018; Nuinoon *et al*., 2022; Ainoon *et al*., 2003; Matsuoka *et al*., 2007; Hsia *et al*., 1993; Bancone *et al*., 2019
21	6	Taipei Chinese-3	Asn→Asp	165	II	China, Phillipines	Tang *et al*., 1992; Hsia *et al*., 1993
22	6	Nankang	Phe→Leu	173	II	China, Malaysia, Vietnam, Singapore	Chen *et al*., 1996; Ainoon *et al*., 2004; Hamada *et al*., 2010
23	6	Miaoli	Phe→Leu	173	II	China	Chen *et al*., 1997
24	6	Coimbra	Arg→Cys	198	II	Myanmar, Indonesia, Malaysia, Vietnam, Cambodia	Matsuoka *et al*., 2004; Sulistyaningrum *et al*., 2020; Ainoon *et al*., 2003; Ngo *et al*., 2022

25	7	Nanning	Leu→Phe	235	III	China	Yan *et al*., 2006
26	8	Bangkok	Lys→Asn	275	I	China	Tanphaichitr *et al*., 2011
27	8	Chinese-1	Thr→Ser	279	II	China	Beutler *et al*., 1992
28	8	Haikou	Thr→Ala	279	II	China	Cai *et al*., 2001
29	8	Baju Maumere	Asp→Tyr	282	III	Indonesia	Kawamotoo *et al*., 2006
30	9	Viangchan / Jammu	Val→Met	291	II	China, Myanmar, Thailand, Malaysia, Vietnam, Laos, Cambodia, Phillipines	Beutler *et al*., 1991; Poon *et al*., 1988; Lee *et al*., 2018; Nuinoon *et al*., 2022; Ainoon *et al*., 2003; Matsuoka *et al*., 2007; Sanephonasa *et al*., 2021; Matsuoka *et al*., 2005; Hsia *et al*., 1993
31	9	Seoul	Gly→Ser	306	II	China	Chao *et al*., 1991
32	9	Kalyan-Kerala	Glu→Lys	317	III	Thailand	Ninokata *et al*., 2006
33	9	Chatham	Ala→Thr	335	II	Indonesia, Malaysia, Phillipines	Hutagalung *et al.*, 2019; Ainoon *et al*., 2003; Hsia *et al*., 1993
34	9	Fushan	Ala→Asp	335	II	China	Xu *et al*., 1995

35	9	Chinese-5	Leu→Phe	342	III	China, Vietnam, Singapore, Cambodia	Chiu *et al*., 1993; Hue *et al*., 2009; Hamada *et al*., 2010; Tantular *et al*., 2021
36	11	Surabaya	Val→Met	431	II	Indonesia	Iwai *et al*., 2001
37	11	*Name unknown*	Tyr→Tyr	437	*NR*	Vietnam	Hue *et al*., 2009
38	11	Andalus	Arg→His	454	II	Malaysia	Ainoon *et al*., 2003
39	11	Union / Chinese-2	Arg→Cys	454	II	China, Myanmar, Malaysia, Vietnam, Phillipines, Cambodia, Laos	Wagner *et al*., 1996; Lee *et al*., 2018; Ainoon *et al*., 2003; Matsuoka *et al*., 2007; Hsia *et al*., 1993; Tantular *et al*., 2021; Sanephonasa *et al*., 2021
40	12	Canton	Arg→Leu	459	II	Myanmar, Thailand, Malaysia, Vietnam, Cambodia, Laos	Lee *et al*., 2018; Nuinoon *et al*., 2022; Ainoon *et al*., 2003; Ngo *et al*., 2022; Tantular *et al*., 2021; Sanephonasa *et al*., 2021

41	12	Yunan	Ala→Thr	461	*NR*	China	Ren *et al*., 2001
42	12	Kaiping/ Anant/ Dhon/ Sapporo-like/ Wosera	Arg→His	463	II	China, Indonesia, Thailand, Malaysia, Vietnam, Phillipines, Laos	Wagner *et al*., 1996; Chiu *et al*., 1991; Sulistyaningrum *et al*., 2020; Nuinoon *et al*., 2022; Ainoon *et al*., 2003; Matsuoka *et al*., 2007; Hsia *et al*., 1993; Sanephonasa *et al*., 2021
43	12	Laibin	Ile→Leu	472	*NR*	China	Yan *et al*., 2006
44	10	Georgia	Tyr→End	428	I	Phillipines	Xu *et al*., 1995
45	4 9	Hechi	Val→Met Val→Met	68 291	II	China, Phillipines	Yan *et al*., 2006; Hsia *et al*., 1993
46	12	Bangkok Noi	Arg→Leu Phe→Cys	459 501	I	Thailand	Tanphaichitr *et al*., 2011

NR- Not reported

2.3.3. Glucose-6-Phosphate Dehydrogenase variants reported in Indian population

The Indian population is socially classified based on strictly defined endogamous castes, tribes and religious groups. Highest population of tribal is found in India compared to other regions of the world (Murthy, 2011). The tribal people constitute about 8.6% of the total population of India (Census of India, 2011). Scheduled tribe and scheduled caste populations have reportedly been shown to have a greater prevalence of G6PD deficiency

(Rai & Kumar, 2012). Although the molecular studies to identify the variants prevalent in different parts of the country are scanty, based on the available literature we have obtained a total of 27 variants that have been identified in Indians both living in India and abroad (Table 7).

Table 7. G6PD variants and mutations identified among the Indians.

Sl. No.	**Variant**	**Mutation**	**Codon**	**Exon No.**	**Class**	**References**
1	Andhra Pradesh	*NR*	*NR*	*NR*	*NR*	Rattazzi, 1968
2	West Bengal	*NR*	*NR*	*NR*	*NR*	Azevedo *et al*., 1968
3	Porbandar	*NR*	*NR*	*NR*	*NR*	Cayanis *et al*., 1977
4	Cutch	*NR*	*NR*	*NR*	*NR*	Goshar, 1979
5	Punjab	*NR*	*NR*	*NR*	*NR*	Verma *et al*., 1987
6	Orissa	Ala→Gly	44	3	III	Kaeda *et al*., 1995
7	Namoru	Tyr→His	70	4	II	Chalvam *et al*., 2007
8	A^{-202}	Val→Met	68	4	III	Rajkhowa *et al*., 2020
9	Hammersmith	Val→Glu	108	5	III	Kaeda *et al*., 1995
10	A^{+}	Asn→Asp	126	5	III-IV	Rajkhowa *et al*., 2020
11	Vanua Luva	Leu→Pro	128	5	II	Kaeda *et al*., 1995
12	Gond	Met→Ile	159	5	*NR*	Sarkar *et al*., 2010
13	Mahidol	Gly→Ser	163	6	III	Bharti *et al*., 2019
14	Madhya Pradesh	Arg→Gly	182	6	*NR*	Devendra *et al*., 2020
15	Mediterranean	Ser→Phe	188	6	II	Beutler *et al*.,1991
16	Coimbra	Arg→Cys	198	6	II	Chalvam *et al*.,

						2008
17	Nilgiri	Arg→His	198	6	II	Chalvam *et al*., 2008
18	Acores	Ile→Val	199	6	*NR*	Bharti *et al*., 2019
19	Jammu	Val→Met	291	9	II	Beutler *et al*., 1991
20	Ludhiana	Gly→Glu	310	9	II	Sukumar *et al*., 2005
21	Kalyan-Kerala / Jamnagar / Rohini	Glu→Lys	317	9	III	Ishwad & Naik, 1984; Ahluwalia *et al*., 1992 Sukumar *et al*., 2005
22	Bhavnagar	Glu→Val	321	9	I	Devendra *et al*., 2019
23	Insuli	Arg→His	330	9	IV	Sukumar *et al*., 2003
24	Chatham	Ala→Thr	335	9	II	Vulliamy *et al*., 1988
25	Guadalajara	Arg→Cys	387	10	I	Vulliamy *et al*., 1998
26	Telti	Leu→Phe	440	11	I	Minucci *et al*., 2012
27	Kaiping	Arg→His	463	12	II	Wang *et al*., 2008

NR- Not reported

Almost 90% of G6PD deficiency in India is due to three common mutations namely, G6PD Mediterranean, G6PD Orissa, and G6PD Kalyan–Kerala (Sukumar *et al*., 2004). One of the common variant among the tribal population of India is the G6PD Orissa. This variant is not seen in the urban population, instead, the Mediterranean variant accounts for most of the G6PD deficiency in the urban population. In the Orissa variant, the stability

of the enzyme is surprisingly increased, while other variants show a reduced stability (Kaeda *et al*., 1995). The variant G6PD Kalyan, identified in a tribal group called Koli living near Mumbai as well as Indians living in North America was later seen in Indians living in Western London (Ishwad & Naik, 1984). This variant was similar to a previously identified variant called G6PD Kerala. Both these variants resulted from a substitution of Glutamic acid by Lysine at amino acid 317. Since both the variants were a result of same mutation, this variant was later termed as G6PD Kalyan-Kerala (Ahluwalia *et al*., 1992). Two other variants namely, Jamnagar and Rohini were also reported to be a result of the same mutation that causes G6PD Kalyan-Kerala. G6PD Jamnagar was reported from Gujarat in a male of 54 years old with haemolysis after ingestion of chloroquine. Whereas, the variant Rohini was reported in a boy of 15 years old from Maharashtra without any prior incidences of haemolysis or neonatal jaundice (Sukumar *et al*., 2005).

A different variant called G6PD Punjab was identified in a Sikh male of Punjabi origin, which was closely related to Kerala, Porbandar and Kalyan variants (Verma *et al*., 1987). In addition to the polymorphic variants, incidences of variants like Hammersmith (108 Val→Glu) and Vanua Luva (128 Leu→Pro) which are much rarer, has also been seen in Indian population living in the other countries. Both these mutations are responsible for lowering the G6PD enzyme activity (Kaeda *et al*., 1995).

Chalvam *et al*. (2007) identified two different variants, Namoru and Nilgiri, among the tribal groups of Nilgiri district of Tamil Nadu. The variant Namoru is caused by an amino acid change at 70 Tyr → His whereas Nilgiri is a result of Arg → His substitution at codon 198. Recent studies have confirmed that this variant is the most commonly occurring variant in Southern India and suggests its origin from India. Another variant called Coimbra (198 Arg→Cys) was also reported in Indian population (Chalvam *et al*., 2007; Chalvam *et al*., 2008). This variant has been categorized as class II variant which is responsible for severe haemolytic anaemia. G6PD Chatham, a mutation resulting from Ala →Thr substitution at amino acid 33, was identified in an Indian boy living in England and in a man from Syria (Vulliamy *et al*., 1988). This variant was first discovered in Northern Iran among the Mazandaranians. This variant is polymorphic and classified as class II group of

variants (Mesbah-Namin *et al*., 2002). Another report of a variant called Kaiping that has not yet been reported from mainland India was discovered in an Indian living in Malaysia. This is a common variant in the Chinese, Thai, Vietnamese and Indonesian population (Wang *et al*., 2008).

In the Northeastern region of India, limited studies have been conducted on G6PD variants. G6PD Mahidol, Acores, A^{+}, A^{-202} are the variants reported from Northeastern Indian population (Bharti *et al*., 2019; Rajkhowa *et al*., 2020). The variant Mahidol is widespread throughout Southeast Asia, with highest prevalence in Thailand and China-Myanmar border (Phompradit *et al*., 2011; Li *et al*., 2015). G6PD Acores is a rare variant which was reported earlier from Portuguese populations and Thailand-Myanmar border populations (Manco *et al*., 2007; Bancone *et al*., 2014). These two variants were observed among the population living in four malaria-endemic states of Northeast India, viz., Meghalaya, Arunachal Pradesh, Mizoram and Tripura (Bharti *et al*., 2019).

Till date, all mutations reported from Indian population are exonic, no intronic mutations have been reported so far. Other variants like Andhra Pradesh, West Bengal, Porbander and Cutch have been reported from Indian populations but the mutations responsible for these variants have not been identified yet (Rattazzi 1968; Azevedo *et al*., 1968; Cayanis *et al*., 1977; Sukumar *et al*., 2003).

2.3.4. Types of mutations involved in various Glucose-6-Phosphate Dehydrogenase deficiency

The types of mutation that are involved in G6PD deficiency are summarized in Table 8. Approximately, 86.08% of all mutations identified till date are caused by single amino acid replacement mutations. The other types of mutations are two or three amino acid replacements, in-frame deletions, nonsense mutation and splicing site mutations (Luzzatto *et al*., 2020).

Table 8. Types of mutations involved in G6PD deficiency (Luzzatto *et al*., 2020).

Type of mutation	Class I	Class II & III	Class IV	Class not defined	Total
Single amino acid replacement	82	87	05	24	198
Two amino acid replacement	4	11	0	0	15
Three amino acid replacement	2	1	0	0	3
In-frame deletions	11	0	0	0	11
Nonsense mutation	0	0	0	1	1
Splicing site mutation	2	0	0	0	2

2.3.5. Distribution of mutations in Glucose-6-Phosphate Dehydrogenase gene

The different types of mutations affecting the exonic regions of G6PD have been summarized exon-wise in Table 9. Highest number of mutations has been reported in exon 10.

Table 9. Location of mutations reported in *g6pd* gene.

Sl. No.	Exon no.	No. of mutation
1	Exon 2	10
2	Exon 3	04
3	Exon 4	19
4	Exon 5	31
5	Exon 6	27
6	Exon 7	19
7	Exon 8	17
8	Exon 9	25

9	Exon 10	41
10	Exon 11	12
11	Exon 12	14
12	Exon 13	05
13	Intron	02

2.4. Haemoglobinopathies

Haemoglobinopathies are the most widespread single-gene disorder worldwide. Approximately 270 million people of the world are affected with haemoglobinopathies, which constitute about 4.5% of the total population of the world (Weatherall & Clegg, 2001). Among all forms of haemoglobinopathies, the thalassaemia major, HbS syndromes, combination of HbE-β-thalassaemia and α-thalassaemia syndromes require special consideration. These syndromes are largely population specific and thus variations occur in the clinical significance of these forms of haemoglobinopathies among populations (Harteveld *et al*., 2022). In the Mediterranean, Middle and Far East, β-thalassaemia major is more common and causes a major health issue. Carrier screening programs are adopted in those regions to control the birth of affected children (Kyriakou & Skordis, 2015). The combined Hb E-β-thalassaemia and α-thalassaemia syndromes are more widespread in the Southeast Asian region (Weatherall, 2016), while the HbS syndromes occur more commonly in the Middle East, Sub-Saharan Africa and India (Weatherall *et al*., 2006).

2.4.1. Prevalence of haemoglobinopathies in India

Haemoglobinopathies and thalassaemias constitute a major health concern in India. In some populations of India, the prevalence rate of β-thalassaemia has reached as high as 17% (Vaz *et al*., 2000). The other common Hb variants that are found to be prevalent in different parts of the country are HbS and HbE syndromes, and HbE-β- thalassaemia (Kamle & Chaturvedi, 2000; Rao *et al*., 2010). A cumulative gene frequency of 4.2% of haemoglobinopathies is observed in India. Haemoglobinopathy carriers is found to be more

than 42 million out of total one billion population of India and every year India experiences over 12000 births carrying some form of haemoglobinopathies (Sarnaik, 2005). The tribal population of Andhra Pradesh and Gujarat has recorded a high prevalence of α-thalassaemia ranging between 1-40%. However, it is fortunate enough that the α-thalassaemia genotype α^{+} is not lethal in both homozygous and heterozygous state (Mukherjee *et al*., 1997). HbS gene was frequently seen in Indian population having Dravidian origin, but with migration and inter-mixing between different ethnic groups, it is now observed in different caste groups and regions of the country. This gene is predominant in the central part of India. The states like Orissa, Andhra Pradesh, Madhya Pradesh and Maharashtra experience high prevalence of HbS gene, while it is comparatively lower in the states like Karnataka, Kerala, Uttar Pradesh, Tamil Nadu (Sarnaik, 2005).

The other common Hb variant in Indian population is HbD. This variant is predominant in the states of Uttar Pradesh, Gujarat, Punjab and Jammu & Kashmir. However, this variant remains clinically asymptomatic in both the homozygous and heterozygous condition. HbD becomes clinically significant only when it is co-inherited with Hb S disease. In this condition, severity sickle cell disease increases (Balgir, 1996; Agarwal *et al*., 1981).

A varying frequency of HbE ranging from 0.002 to 0.646 has been reported from different regions of India (Madan *et al*., 2010; Deka *et al*., 1988). This variant is more common in the Northeastern states of India, which in some areas have reached carrier frequencies that is as high as 50% (National Health Mission, 2016). There are several other rare Hb disorders that do not have much significant clinical consequences. These variants include hereditary persistence of foetal haemoglobin (HPFH), Hb Lepore, Hb Chandigarh, HbQ, HbK and many others (Sarnaik, 2005).

2.5. Association of Glucose-6-Phosphate Dehydrogenase deficiency with haemoglobinopathies

It is considered that the structural haemoglobin variants and thalassaemias are protected against malaria by the process of natural selection. Evidences supporting this observation have been found in *P. falciparum* infections via *in-vitro*, epidemiological studies and cartographic modeling experiments (Weatherall, 2008). However, clinical studies only will confirm these associations which are yet to be made (Taylor *et al*., 2012b). Studies have compared the incidence of G6PD deficiency in sickle cell patients with general population. However, many but not all studies confirmed this association (Bouanga *et al*., 1998). In malaria endemic regions, both sickle cell anaemia and G6PD deficiency provide some advantage over the normal population. Individuals having either of these conditions have less chance of being infected with malaria (Nkhoma *et al*., 2009).

2.6. Clinical manifestations of Glucose-6-Phosphate Dehydrogenase deficiency

2.6.1. Neonatal jaundice, hyperbilirubinemia and kernicterus

G6PD deficiency becomes a major risk factor that may result from haemolysis (Garcia *et al*., 2021) in neonatal jaundice. Elevated production and reduced elimination of bilirubin results in increased serum bilirubin level causing hyperbilirubinemia. This high bilirubin is capable of passing the blood-brain barrier and cause neurological conditions like bilirubin encephalopathy or kernicterus (Kaplan *et al*., 2004). However, all G6PD deficient neonates do not develop neonatal jaundice. The risk of developing jaundice and its severity depends on the variants and at times on other co-existing factors (Lee *et al*., 2022). Significant intensity of hyperbilirubinemia entailing phototherapy was observed in the variant Kaiping (Wong *et al*., 2017). Other variants from Southeast Asian region that were observed to be associated with hyperbilirubinemia are Mediterranean (Wong *et al*., 2017; Ezz El-Deen *et al*., 2013), Sapporo-like, Canton, Gaohe, Viangchan (Wong *et al*., 2017), $A^{-}{}^{202}$, Jammu, Chinese-2, Union (Vela-Amieva *et al*., 2021), Mahidol (Wong *et al*., 2017; Vela-Amieva *et al*., 2021), Orissa (Moiz *et al*., 2012).

2.6.2. Acute haemolytic anaemia

Among all triggers of haemolysis in G6PD deficient persons, consumption of fava beans is the most common trigger, which causes a brisk haemolysis. The second type of haemolysis is drug induced haemolysis. Both of these factors trigger a dose-dependent haemolysis in G6PD deficient people. In many cases, haemolysis may be compensated and may not cause anaemia. Again, in case of mild anaemia it may not be necessary to visit the health care centres, as a result of which the condition remains undiagnosed. The third potential trigger of acute haemolytic anaemia in G6PD deficiency is infection. This type of haemolysis, however, is very unpredictable. In order to result in acute haemolytic anaemia by bacterial infection, the infection must be very severe. Such incidences of acute haemolytic anaemia has been reported with Maxillary abscesses caused by *Streptococcus* (Quereshy *et al*., 2000) or *Staphylococcus* (Paek *et al*., 2009) and *Clostridium difficile* infection (Lodha *et al*., 2008) . Acute haemolytic anaemia has also been reported in Hepatitis A, B (Chau *et al*., 1997) and E (Abid *et al*., 2002; Zamvar *et al*., 2005). There is also an instance of acute haemolytic anaemia in dengue fever (Khan *et al*., 2019).

2.6.3. Chronic nonspherocytic haemolytic anaemia

The risk of acute haemolytic anaemia in G6PD deficiency is a common complication; while, chronic nonspherocytic haemolytic anaemia (CNSHA) in G6PD deficiency is a rare condition (Tanphaichitr *et al*., 2011), with an estimated frequency of <10 per million. Similar to Hereditary Spherocytosis, the clinical picture in CNSHA might be moderate (found inadvertently in adults) or severe enough (found in a small percentage of patients) to necessitate repeated blood transfusions. Unlike in Hereditary Spherocytosis, the typical presentation is a history of severe neonatal jaundice. Patients with CNSHA will experience either acute or chronic haemolysis from any agents that can cause acute hemolytic anaemia in a G6PD deficient individual (Luzzatto *et al*., 2020).

2.7. Approaches attempted for treatment of Glucose-6-Phosphate Dehydrogenase deficiency

Although G6PD-associated enzymopathy was discovered more than 60 years ago, the belief that common variants have a mild pathology that can be overcome by avoiding trigger foods and drugs or treated with blood transfusions during a haemolytic crisis hinders the development of a therapeutic measure (Garcia *et al*., 2021). However, there are a few intervention studies which were attempted to overcome G6PD deficiency associated challenges.

2.7.1. Antioxidant therapy

Four potent antioxidants have been studied to address G6PD deficiency, the first of which is ascorbic acid, or commonly referred to as Vitamin C. An *in-vivo* study on ascorbic acid have shown it to have a pro-oxidant effect on the RBCs and use of high dose ascorbic acid was seen to induce acute haemolytic anaemia in G6PD deficient patients (Pearson *et al*., 2021; Fujii *et al*., 2019). Again, a lower dose of ascorbic acid was used to take care of methaemoglobinaemia in the G6PD deficient patients. This was successful in treating methaemoglobinaemia, however blood transfusions were required as co-treatments, thereby veiling up the efficacy of low dose ascorbic acid treatment (Reeves *et al*., 2016; Sonbol *et al*., 2013; Rehman *et al*., 2018).

α-Tocopherol, commonly called Vitamin E, is an antioxidant that has the potential to protect the cell membrane from oxidative damage. Several case studies have been reported on the use of α-Tocopherol to treat G6PD deficiency. While some studies have shown that the hematological parameters have improved upon treatment with α-Tocopherol, but others were of opposite opinion (Spielberg, 1979; Hafez *et al*., 1986; Johnson *et al*., 1983; Newman *et al*., 1979). Again, although improvements in the haematological parameters were observed, but when the patients RBCs were tested with H_2O_2 *in vitro*, no changes were observed in the degree of haemolysis (Corash *et al*., 1980).

Carotenoid (astaxanthin) was seen to boost the activity of purified wild type (WT) G6PD in a dose-dependent manner (Temel *et al*., 2017). Other carotenoids showed antioxidant effect *in vivo* but did not protect against glutathione (GSH) reduction (Chiste *et al*., 2014; Koriem & Arbid, 2018).

α-Lipoic acid helped the total antioxidant capacity of the erythrocytes in both WT and deficient G6PD. However, it does not show significant effect in case of an acute oxidative stressor (Georgakouli *et al*., 2013; Georgakouli *et al*., 2018).

2.7.2. Enhancement of Glutathione (GSH) biosynthesis via N-acetyl-cysteine (NAC)

In G6PD deficient patients, regeneration of GSH via NADPH-dependent pathway is the only way to restore GSH pool (Raftos *et al*., 2010). NAC, a prodrug has been reported to improve plasma L-cysteine (LC) which increases GSH levels in the erythrocytes (Pei *et al*., 2018). The plasma LC generated via NAC supplementation has the capability of increasing the erythrocytes' GSH biosynthesis *in vitro* (Fazary *et al*., 2020; Raftos *et al*., 2007).

2.7.3. Nicotinamide Adenine Dinucleotide Phosphate (NADPH) generation via complementary pathways

There are three metabolic ways through which NADPH is produced in the cytosol. The first route is the PPP, in which G6PD employs glucose to produce NADPH, the other two ways are through the Tricarboxylic acid (TCA) cycle where Malic Enzyme 1 (ME1) and Isocitrate Dehydrogenase 1 (IDH1) are involved. ME1 and IDH1 utilize the TCA cycle intermediates for generation of NADPH and are dependent on the mitochondrial redox balance (Xiao *et al*., 2018). In the erythrocytes, due to the absence of mitochondria, NADPH production is mainly dependent on G6PD. However, even though mitochondrion is absent, they still express ME1 and IDH1 along with the TCA cycle enzymes (D'Alessandro *et al*., 2017). Again, in the G6PD deficient erythrocytes, precursors of these complementary routes are found which are capable of generating NADPH by utilizing the TCA cycle intermediates (Francis *et al*., 2020). But, this NADPH is not sufficient enough to tackle against the oxidative stress in the G6PD deficient erythrocytes. It is yet to be confirmed if augmenting these complementary pathways in G6PD deficient erythrocytes

would be adequate enough to tackle against the high oxidative stress and resultant haemolysis (Garcia *et al*., 2021).

2.7.4. Using transcriptional regulators to target upregulation of Glucose-6-phosphate dehydrogenase

During the process of RBC production (erythropoesis), the erythroblasts go through sequential changes including nuclear extrusion among others (Menon and Ghaffari, 2021). Even though this process is necessary for functioning of the erythrocytes, however, it poses a therapeutic challenge since the RBCs are unable to produce new proteins. Targeting transcription in the early phase of erythroblasts before the process of nuclear extrusion can be a smart approach to boost G6PD protein levels. This elevated production of G6PD can be helpful in treatment of G6PD deficiency associated conditions (Garcia *et al*., 2021).

2.7.5. Small molecule activators to address Glucose-6-Phosphate Dehydrogenase deficiency

A structural chaperone, AG1 was discovered recently which was capable of partially correcting the catalytic activity and stability of several class II and III variants *in-vitro*, *in-cellulo* and *in-vivo* as well. The small molecule, AG1 was successful in enhancing NADPH and GSH levels, and reducing ROS, thereby protecting the cells against damage caused by oxidative stress agents (Hwang *et al.,* 2018). Again, a structure-based pharmacophore screening using AG1 as a scaffold identified five compounds as potent activators of G6PD (Saddala *et al*., 2020). The recent development in the structure of the most commonly occurring class I variants have shown that these variants have a common structural fault. This has created an opportunity for developing a common structural chaperone that can be attempted for correcting this common fault (Horikoshi *et al*., 2021).

Owing to the fact that Class II and III variants although are common but often asymptomatic, and the occurrence of more severe forms of the deficiency is not common, pharmaceutical firms are not as interested in coming up with a treatment for this condition.

Furthermore, it is unlikely that there will be a "one-size-fits-all" therapy due to the large number of mutations and their varied effects on protein structure. More studies on the disease epidemiology confirming the existing needs for developing a therapy against the deficiency will draw the interest of the pharmaceutical sector.

2.8. *In-silico* drug designing for Glucose-6-Phosphate Dehydrogenase variants

The process of conventional drug discovery starting from identifying the target to the launching of a final product is a complex one. The cost involved is extremely high and the process takes about 12-15 years. Currently, the recent development in computer potency and technology has helped to attain the purpose quite simpler and speed up the entire process through computer-aided drug design (CADD). The *in-silico* methods like molecular docking and molecular dynamics simulation are the vital and frequently used methods in the process of CADD. The development of the well-known drugs such as Indinavir, Sequinavir, Ritonavir, Tirofiban are some of the examples where CADD has been successful.

From the past few decades, molecular docking has turned into a widely used fast and inexpensive technique for drug discovery, both in academic and industrial settings. Virtual screening by means of molecular docking can be grouped into two types on the basis of the approach used: ligand-based docking and structure-based docking. In ligand-based docking, the ligands are known but the structure of the target protein needs to be identified. On the other hand, in structure-based docking, the structure of the target proteins is known and the potential ligands for the specific target should be identified.

Currently, there are a number of protein-ligand docking programs among which AutoDock, DOCK, FlexAID, LeDock, Glide, rDock, SEED are commonly used ones. AutoDock Vina, a suite of automated docking tools developed by The Scripps Research Institute is the most commonly used open source docking programs among other docking programs of the AutoDock suite as well as other programs.

Molecular dynamics (MD) simulation has gained wide application in the field of modern drug discovery. MD simulation is a computer simulation technique used to obtain structural, dynamical and thermodynamical details of a molecular system such as proteins, enzymes or lipids (Badar *et al*., 2022).

There are only a few studies of G6PD focusing on the protein-ligand affinities and structural integrity for G6PD monomers and dimers (Doss *et al*., 2016. These studies demonstrated the utility of combining structural and computational research to comprehend the disease's genotype-phenotype association.