Glucose 6-Phosphate Dehydrogenase Mutations Causing Enzyme Deficiency in a Model of the Tertiary Structure of the Human Enzyme


Human glucose 6-phosphate dehydrogenase (G6PD) has a particularly large number of variants resulting from point mutations; some 60 mutations have been sequenced to date. Many variants, some polymorphic, are associated with enzyme deficiency. Certain variants have severe clinical manifestations; for such variants, the mutant enzyme almost always displays a reduced thermal stability. A homology model of human G6PD has been built, based on the three-dimensional structure of the enzyme from Leuconostoc mesenteroides. The model has suggested structural reasons for the diminished enzyme stability and hence for deficiency. It has shown that a cluster of mutations in exon 10, resulting in severe clinical symptoms, occurs at or near the dimer interface of the enzyme, that the eight-residue deletion in the variant Nara is at a surface loop, and that the two mutations in the A-variant are close together in the three-dimensional structure.

© 1996 by The American Society of Hematology.

GLUCOSE-6-PHOSPHATE dehydrogenase (G6PD) is a housekeeping enzyme that catalyzes the first and rate-limiting step in the pentose phosphate pathway. Its key role in metabolism is to provide reducing power in the form of NADPH. This role is particularly important in red blood cells (RBCs) where NADPH is required for detoxification of hydrogen peroxide and other compounds, via glutathione. In these cells the reaction catalyzed by G6PD (and by 6-phosphogluconate dehydrogenase, which depends on G6PD) is the only source of NADPH. Human G6PD, a homodimer encoded by a gene which maps to the region Xq28 on the X-chromosome, exhibits an extraordinary degree of genetic variability: more than 100 deficient variants have been reported and characterized. These variants are either true polymorphisms, with relatively mild clinical manifestations and variant alleles reaching frequencies of 1% to 50% in various parts of the world, or sporadic variants, brought to light because they cause chronic nonspherocytic hemolytic anemia (CNSHA) in affected males. The former group of G6PD-deficient variants have most likely reached polymorphic frequencies because deficient individuals are protected to some extent against severe Plasmodium falciparum malaria. In both polymorphic and sporadic variants there is always some residual enzyme activity and this is invariably lower in RBCs than in other cells. The lack of null variants and the low activity in RBCs (which cannot make up for enzyme breakdown through de novo protein synthesis) implies that instability of mutant G6PD molecules is probably the commonest cause of G6PD deficiency. This has been confirmed in several cases by appropriate biochemical analysis.

Following the determination of the primary structure of human G6PD through cloning of the corresponding cDNA, the molecular lesions of more than 60 G6PD variants have been determined, nearly all are missense point mutations causing single amino acid replacements. These mutations, with differing clinical manifestations, vary in their effect on residual enzyme activity and on substrate binding; almost all cause a decrease in enzyme stability. Figure 1 shows the distribution of severe (class I) mutations with respect to the primary structure of the protein; notable is a cluster of severe mutations between residues 380 and 450. This analysis has proved most valuable for diagnosis and to define which mutations account for G6PD deficiency in various populations. However, it has been less successful in defining the role of different domains and of individual amino acid residues in the stability of the protein, and in explaining why some mutants are almost asymptomatic whereas others cause severe CNSHA.

Some of us have recently solved and given a detailed description of the three-dimensional structure of G6PD from Leuconostoc mesenteroides. The protein from this species has significant homology to human G6PD, and residues shown to be important for activity are conserved. Thus, we can use knowledge of the structure of the bacterial enzyme to model normal human G6PD B. The dimer is shown to be the smallest molecular form of G6PD that is active in the human enzyme as in all species. The crystal structure of the L. mesenteroides apo-enzyme has a dimer in the asymmetric unit. The two monomers are related by an axis which is very close to a dyad and it is anticipated that the subunit interface is retained in the human enzyme. Our homology model allows us to predict, with some confidence, the position and role of both normal and mutated residues in the tertiary structure of human G6PD, and thus answer some of these questions.

MATERIALS AND METHODS

The human model was generated by homology modeling from the known structure of L. mesenteroides G6PD, the 33% sequence homology...
Fig 1. Sequence alignment of human and \textit{L mesenteroides} GGPD. The alignment was used for model building of human GGPD. Residues fully or highly conserved over 14 sequences are indicated in different shades of gray. The secondary structural elements for the \textit{L mesenteroides} enzyme are shown. The locations of amino acid substitutions found in severe (class I) variants are indicated by capital letters above the alignment. Note the cluster found around amino acid 400.
identity is sufficient to justify the assumption that it has a similar fold. A multiple sequence alignment of 14 currently known G6PD sequences (human, L mesenteroides, Synechococcus PCC 7942, Zymomonas mobilis, Saccharomyces cerevisiae, Drosophila melanogaster, rat, mouse, Erwinia chrysanthemii [N. Hegouvieux-Cotte-Pattat, J. Robert-Baudouy, PIR database, entry S37053], Esherichia coli, Kluyveromyces marxianus lactis [M. Wesolowski-Louvel, C. Tanguy-Rougeau, H. Fukuhau, PIR database, entry S31337], Plasmodium falciparum, Pichia jadinii, Bacillus SP HT-3H. [H Sagai, K Hattori, M Takahashi, US Patent No. 5,137,821,
A 3D MODEL OF DEFICIENT MUTANTS IN HUMAN G6PD

Fig 2. Model of the human dimer. In this figure subunit P (equivalent to subunit A in published illustrations of the dimer of L. mesenteroides G6PD) is shown in red, green, and yellow, and Q is in pale blue, pink, and white. The monomer consists of two domains—a smaller coenzyme domain encompassing residues 1-198 and a larger β + α domain comprising residues 198-515. The sequence GAGSYAQ (residues 38-44) is at the coenzyme binding site (arrow). The G6P binding site includes residues from the perfectly conserved 9-amino acid sequence RIDHYLGKE domain encompassing residues 1-198 and a larger region used for this and other figures is shown schematically at upper right as a guide for the reader. Figures 2, 4, 5, and 6 were drawn using the program BOBSCRIPT, an enhanced version of MOLSCRIPT.41

Fig 4. Serious mutations at the dimer interface. A magnification of the dimer interface region of the model. Subunit P is in red, green, and yellow, Q in pale blue, pink, and white. The Ca carbons of mutations causing seriously deficient (class I) variants are indicated by black spheres in subunit P and grey spheres in Q. All variants are labeled and the region of the dimer enlarged is shown at upper right. The class I variant residues form two symmetry-related clusters at the dimer interface. The cluster at the lower right hand corner of the diagram involves residues from subunit P: 213Val → Leu (Minnesota3), 216Phe → Leu (Harlaud4) (not visible), 274Glu → Lys (Coom35), 278Ser → Phe ("Wesham"44), 393Arg → His (Nashville26), 394Val → Lys (Alhambra27), 396Pro → Lys (Bari7) and 398Glu → Lys (Puerto Limon28); and from subunit Q: 386Cys → Arg (Tomah49), 398Lys → Glu (Iowa49), 397Arg → Cys (Guadalajara16) 387Arg → His (Beverly Hills40, 405Met → Ile (Clinic40), 410Gly → Cys (Riverside37); 410Gly → Asp (Japan41) and 416Glu → Lys (Tokyo42). The other cluster contains the same residues in the alternative subunit. Further from the cluster are 363Asn → Lys (Loma Linda27), and 439Arg → Pro (Pawnee33). These residues interact across the dimer interface, generating the major interaction energy between the two subunits. Mutations disrupt and weaken the dimer-dimer interactions, leading to destabilization. Highlighted in the diagram is Lys 386, originally predicted as a residue interacting with NADP+.5

Fig 5. The eight-residue deletion situated in a flexible loop. A magnification of the region of the eight-residue deletion in the class I variant, Nara, which is predicted to be missing residues 318-325. This is the only variant identified so far that has more than two residues deleted. The first and last deleted residues are indicated by black spheres at their Ca positions. The chain trace highlighted in red (A) shows the modeled conformation of the normal enzyme at this point. The blue chain trace in (B) shows a suggested alternative conformation after deletion of residues 318-325. This alternative conformation was created by removing residues 318 to 325 and then modeling a new conformation using Ca surrounding residues were altered as little as possible. No energy minimization was performed. It is possible to model the deletion without affecting any secondary structural elements or the active site of the protein.

Fig 6. The region including 68Val → Met and 126Asn → Asp of the A- deficient variant. A magnification of helix αc and the β-sheet of the coenzyme binding domain. The residues mutated in the variant A- are shown in ball and stick (grey, carbon; blue, nitrogen; red, oxygen; yellow, sulphur) to indicate their proximity. The left-hand diagram shows the residues in the common G6PD A- (Met 68 and Asp 126). Val 68, on βB, is also a valine in L. mesenteroides G6PD; the sheet strand is immediately followed by the conserved Arg 72 and the position of the sidechain of Val 68 can thus be modeled with confidence. The common African variant A (126Asn → Asp) is not deficient nor is the engineered mutant 68Val → Met. The normal behavior of the mutant 68Val → Met allows us to assume that βB remains unchanged and that the sidechain of Met 68 will be on the same side of the sheet as is Val 68. In the model, the sidechains of Val 68 and Met 125 are in direct contact; direct contact between 68 and 126 may be achieved by a small rotation of αc on its axis.

1992)) was generated, using Alignment of Multiple Protein Sequences (AMPS).9 The resulting alignment is shown for the human and L mesenteroides enzymes (Fig 1). The secondary structural elements for the bacterial enzyme, also given in Fig 1, were identified from its refined three-dimensional structure using the program DSSP29 as is described in Rowland et al.4 This alignment was used with the program MUTATE (R. Einstouf, D. Phil. thesis, Oxford University, 1992) to exchange the sidechains of L mesenteroides G6PD with those in the human enzyme and place each altered sidechain at a plausible main chain conformation was chosen that avoided strain on its axis. Total energy decreased from 265,116 kcal mol⁻¹ to −7,649 kcal mol⁻¹.

RESULTS

The use of 14 sequences and a multiple alignment algorithm such as that in AMPS enhances the reliability of the alignment over that from a pairwise comparison of the two sequences of interest. The method uses a matrix of pairwise identities and optimizes the conservation between all pairs of sequences; thus, it can take advantage of subgroups with higher conservation in comparing two more distantly related sequences. Although the human sequence has 33.1% identity with that from L mesenteroides, uncertainties can be resolved by using, for instance, the 48.1% identity of human G6PD with that of S cerevisiae, which itself has 35.7% identity with L mesenteroides G6PD. The alignment in Fig 1 was achieved without disruption of secondary structure except for a shortening of helices αb and of αk by one turn and a lengthening of αb' to the same extent. All three helices are at the extreme edges of the dimer. The multiple sequence alignment has resulted in several regions where there are one or two residue insertions and deletions, usually in loops. The position of these discontinuities results from optimizing sequence similarities over the complete set of sequences;
one-to-one correspondence of residues cannot be assumed in these regions. Most insertions and deletions arise from preserving the alignment of the *L. mesenteroides* sequence with those of other prokaryotes and the human sequence with the eukaryotic sequences. Typical examples are the loop between βK and βL, whose length varies from 14 to 19 residues, and the insertion/deletion at αβ-αβ' where the exact position of the corner between the two external helices may vary between the two groups of structures. After residue 497 there is considerable variation between sequences and the alignment is rather arbitrary. Although the βM-βN loop is lengthened by one residue in two of the bacterial sequences and by three residues in that of *Z. mobilis*, the structural elements which make up the dimer surface, αf, αg, βM and βN, are not subject to insertions or deletions.

The model shows the extended nature of the dimer (Fig 2) with “β + α” domains forming the dimer interface and “coenzyme” domains distant from the dyad axis. The following description of the human enzyme is based on the detailed description given of *L. mesenteroides* G6PD. The domain structure as defined for the bacterial enzyme is retained; the domain boundary is at the end of sheet strand βF. This is at the terminus of a standard dinucleotide binding fold and is in a region of high sequence conservation. The N-terminal dinucleotide binding domain comprises residues 27-200 in the human enzyme (1-177 of the *L. mesenteroides* enzyme). The conserved NAD(P) fingerprint forms a tight turn that begins at Gly 38; its position is indicated in the figure. We have shown this to be the NADP⁺ binding site in *L. mesenteroides* G6PD (C.E.N., unpublished results, November 1994). The coenzyme specificity is ensured by the presence of the conserved Arg 72, which binds the 2'-phosphate of NADP. The active site can be identified at the domain boundary and includes residues from a nonapeptide (198-206), which is conserved in the 14 G6PD species considered, as well as residues distant in the primary structure. His 201, which is predicted to interact with the substrate (G6P), is 17 Å from Gly 38, the neighbor of the adenine ribose of NADP⁺.

The dimer interface of human G6PD is formed by association of the sheets and two helices in the second (β + α) domain of each subunit to form a barrel. The residues involved in the dimer surface (defined as those which contain atoms which are less accessible to solvent in the dimer than in the putative monomer) are shown schematically in Fig 3. The two active sites (His 201 of the G6P binding sites) in the dimer are more than 50 Å apart, separated by the large, predominantly antiparallel sheets of the dimer interface which together form a half-barrel.

Some 29 class I variants have been considered; they comprise two deletions, of 2 (Stonybrook) and 8 (Nara) residues, and 26 point mutations. The mutated positions are shown in Fig 1. Both deletions occur in surface loops distant from the dimer axis; they are close to each other. Seventeen of the remaining point mutations (20 variants) are in or close to the dimer surface (Table 1). The variant residue is defined as being a part of the surface if one of the atoms in the corresponding residue in G6PD B has a changed solvent accessibility as discussed above. The class I residues involved in the dimer surface are 213 (Minnesota), 216 (Hari-...
A 3D model of deficient mutants in human G6PD

Table 1. Class I Mutations Close to the Dimer Surface

<table>
<thead>
<tr>
<th>Variant</th>
<th>Mutation</th>
<th>Secondary Structure Element</th>
<th>Distance From Dimer Interface</th>
<th>Distance From Active Site</th>
<th>Subunit Involved in Patch</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minnesota</td>
<td>213 Val → Leu</td>
<td>α'</td>
<td>In surface</td>
<td>&gt;10 Å</td>
<td>P</td>
<td>26</td>
</tr>
<tr>
<td>Harlauc</td>
<td>216 Phe → Leu</td>
<td>αf</td>
<td>In surface</td>
<td>&gt;10 Å</td>
<td>P</td>
<td>27</td>
</tr>
<tr>
<td>Corum</td>
<td>274 Gln → Lys</td>
<td>α1−α2</td>
<td>8.7 Å</td>
<td>&gt;10 Å</td>
<td>P'</td>
<td>28</td>
</tr>
<tr>
<td>&quot;Wexham&quot;</td>
<td>278 Ser → Phe</td>
<td>α1−α4</td>
<td>In surface</td>
<td>&gt;10 Å</td>
<td>P'</td>
<td>28</td>
</tr>
<tr>
<td>Tomah</td>
<td>385 Cys → Arg</td>
<td>βK−βL</td>
<td>1.4 Å</td>
<td>&gt;10 Å</td>
<td>Q</td>
<td>29</td>
</tr>
<tr>
<td>Iowa</td>
<td>386 Lys → Glu</td>
<td>βK−βL</td>
<td>In surface</td>
<td>&gt;10 Å</td>
<td>Q</td>
<td>29</td>
</tr>
<tr>
<td>Guadalajara</td>
<td>387 Arg → Cys</td>
<td>βK−βL</td>
<td>2.2 Å</td>
<td>&gt;10 Å</td>
<td>Q</td>
<td>31</td>
</tr>
<tr>
<td>Beverley Hills</td>
<td>391 Arg → His</td>
<td>βL</td>
<td>4.2 Å</td>
<td>&gt;10 Å</td>
<td>P</td>
<td>26</td>
</tr>
<tr>
<td>Nashville</td>
<td>393 Arg → His</td>
<td>βL</td>
<td>5.2 Å</td>
<td>&gt;10 Å</td>
<td>P</td>
<td>31</td>
</tr>
<tr>
<td>Alhambra</td>
<td>394 Val → Leu</td>
<td>βL</td>
<td>5.2 Å</td>
<td>&gt;10 Å</td>
<td>P</td>
<td>40</td>
</tr>
<tr>
<td>Bari</td>
<td>396 Pro → Leu</td>
<td>βL−βM</td>
<td>8.6 Å</td>
<td>&gt;10 Å</td>
<td>P</td>
<td>33</td>
</tr>
<tr>
<td>Puerto Limon</td>
<td>398 Gln → Lys</td>
<td>βL−βM</td>
<td>4.1 Å</td>
<td>&gt;10 Å</td>
<td>P</td>
<td>33</td>
</tr>
<tr>
<td>Clinic</td>
<td>405 Met → Ile</td>
<td>βM</td>
<td>In surface</td>
<td>&gt;10 Å</td>
<td>Q</td>
<td>30</td>
</tr>
<tr>
<td>Riverside</td>
<td>410 Gly → Cys</td>
<td>βM−βN</td>
<td>In surface</td>
<td>&gt;10 Å</td>
<td>Q</td>
<td>29</td>
</tr>
<tr>
<td>Japan</td>
<td>410 Gly → Asp</td>
<td>αD</td>
<td>4.5 Å</td>
<td>&gt;10 Å</td>
<td>Q</td>
<td>32</td>
</tr>
<tr>
<td>Tokyo</td>
<td>416 Gln → Lys</td>
<td>βN</td>
<td>In surface</td>
<td>&gt;10 Å</td>
<td>Q</td>
<td>30</td>
</tr>
<tr>
<td>Pawnee</td>
<td>439 Arg → Pro</td>
<td>αm</td>
<td>In surface</td>
<td>&gt;10 Å</td>
<td>P</td>
<td>31</td>
</tr>
<tr>
<td>Kobe</td>
<td>440 Leu → Phe</td>
<td>αm</td>
<td>4.8 Å</td>
<td>&gt;5.9 Å</td>
<td>P</td>
<td>34</td>
</tr>
<tr>
<td>Santiago de Cuba</td>
<td>447 Gly → Arg</td>
<td>End αm</td>
<td>2.9 Å</td>
<td>&gt;10 Å</td>
<td>P</td>
<td>35</td>
</tr>
</tbody>
</table>

Based on this model of the human G6PD dimer, we have attempted to rationalize the previously reported pattern of naturally occurring mutations and to analyze some specific examples. The model of the human enzyme has allowed us to locate the different variants on particular secondary structural elements and to see their spatial relationships to the coenzyme binding site, the active site and the dimer interface. Almost all residues seen to be important in binding NADP in L mesenteroides G6PD, and those predicted to bind substrate and to promote catalysis, have been conserved in the human enzyme. None of these residues has been mutated in any of the variants described so far. Coenzyme and substrate binding have been modeled by Rowland et al.; the three-dimensional structure of a binary complex of the L mesenteroides enzyme confirms the features of NADP binding (C.E.N., unpublished results, November 1994) and this will not be elaborated further here.

Regions of the surface of the dimer that may be of importance for enzyme stability may be recognized in that class I variants cluster in such an area. It is observations of this kind which we will now address in this report; it is not our intention to use our homology model to answer detailed questions concerning the particular contacts made between variant residues and their neighbors in the three-dimensional structure. It is clear that a detailed analysis of the different sidechain-sidechain interactions that arise in each variant and their consequences in terms of the stability and activity of the human G6PD molecule must await the three-dimensional structure determination of both human G6PD B and of the different variants. We will confine our detailed discussion to three particular examples.

Class I variants: Subunit contacts. A striking previous observation has been that a set of severely deficient variants associated with CNSHA is clustered in exons 10 and 11, and it has been suggested that they correspond to the NADP+ binding site. It is now clear that the binding site is elsewhere, and this cluster corresponds in fact to the subunit interface. These variants are located close to the ends of the sheet strands of the β + α domain; they form two surfaces, each shared by both subunits (Fig 4). The surfaces also include some severely deficient mutations mapping outside exons 10 and 11. The fold brings residues from different regions of primary structure in the same monomer close together in space: residues 213 and 216 on αf are a part of the same surface patch as 393, 394, 396, and 398 on βL and in the βL−βM turn. Similarly, the dyad axis brings residue 405, at the end of βM, and 410, in the βM−βN turn, of the second subunit into the same surface patch. It can be seen that residues at the two ends of βL contribute to different patches; it is the residues that precede βL of the second subunit, 385, 386, and 387, which are in the same cluster as 393, 394, 396, and 398 of the first subunit (Fig 4). It should be pointed out that, of the class I residues close to or in the dimer surface, only residue 440 is within 10 Å of the modeled substrate site of its own subunit. None of the residues is within 10 Å of either coenzyme site or of the substrate site in the second subunit. The large distance between the two active sites of the dimer has already been noted.

Mutations affect both hydrophobic and charge-charge in-
teractions. Salt bridges may be broken: for instance in the variants G6PD Corum,\(^28\) 274Glu → Lys, and G6PD Iowa,\(^29\) 386Lys → Glu. Glu 274 makes a salt bridge to Arg 348 and Lys 386 is a neighbor of Glu 417; G6PD Iowa is described as having greatly decreased thermal stability (E. Beutler, L. Forman, P.A. Alarcón, unpublished results, 1986). The effect of NADP\(^+\) and NADPH in changing the equilibrium between monomers, dimers, and tetramers\(^26\) together with the possibility of reactivating mutants such as G6PD Iowa in high concentrations of NADP\(^+\) has led to the proposal of Lys 386 as an NADP\(^+\) binding residue.\(^29\) However, it is 36 Å from His 201, thought to be the site of substrate binding and is even more distant from the known coenzyme site of the \(L\) *mesenteroides* enzyme, and so is unlikely to bind catalytic NADP\(^+\). The rate-limiting dimerization stage in the reactivation of urea denatured \(L\) *mesenteroides* G6PD has been shown to be enhanced by addition of coenzyme.\(^37\) Stimulation of these processes by ligands is likely to arise by their stabilizing an important intermediate (or the end product) in the folding pathway by forming a binary complex; for \(L\) *mesenteroides* G6PD an intermediate is indicated. The position of Lys 386 in the model for the human enzyme, in the dimer interface and distant from the proposed NADP\(^+\) site, would suggest a similar mechanism.

Hydrophobic contact surfaces at the interface are decreased in some variants: for instance, in G6PD Harilaou,\(^37\) 216Phe → Leu. Two variants, namely G6PD Clinic,\(^39\) 405Met → Ile, and G6PD Minnesota,\(^28\) 213Val → Leu, disrupt the same hydrophobic contact, which involves residues from exon 7 as well as exon 10. The sidechains of these residues on different monomers make contact across the dimer interface. The large number of class I variants in this part of the dimer contact area suggests that the contribution of these residues to inter-subunit contacts is important for the integrity of the dimer and that this region of the interface is particularly sensitive to change and is crucial to stability of the enzyme.

**Class I variant with an eight-residue deletion.** Only one deletion of more than two residues has been described thus far. In G6PD Nara,\(^22\) associated with CNSHA, residues 318-325 are deleted. Some of these residues are part of a surface loop that was found to be highly mobile in \(L\) *mesenteroides* G6PD. There are insertions and deletions in the sequence alignment in this region (Fig 1). The deletion in the human enzyme at 304 is secure and is a prokaryote/eukaryote difference; the alignment of \(βH\), which is well conserved, generally preserves the amphiphilic character and has equivalent small and large residues. The residues of the loop between 311 and 317 need not be equivalent in the three-dimensional structure, but equivalence is certainly achieved by Tyr 322. In the crystal structure of the bacterial enzyme, the residues that are aligned with those in the deletion immediately follow the only loop for which sidechain conformations are not clearly identifiable. The loop has the highest temperature factors seen in the \(L\) *mesenteroides* G6PD structure;\(^2\) the average main-chain temperature factor for the residues 292-297 (312-317 in human G6PD) is 64.5 Å\(^2\); whereas that for the whole protein is 33.7 Å\(^2\).\(^2\) The great flexibility of this loop in G6PD from \(L\) *mesenteroides* suggests that a loop of different length and conformation as proposed for the normal human enzyme is readily accommodated (Fig 5A). Removal of the eight residues 318-325, including the totally conserved Tyr 322, can be accommodated by a local rearrangement of the main chain (Fig 5B). A direct connection between residues 317 and 326 can be made without disruption of the structure but at the expense of exposing otherwise buried residues. The major loss of activity in the human enzyme is likely to arise from deletion of Tyr 322, which shields these hydrophobic residues, Val 309 and Gly 310, from solvent. This would explain why G6PD Nara still has residual activity but is highly unstable.

**Less deleterious variants: G6PD A-.** Not all sequenced variants have been classified in terms of activity. As well as the class I variants, considered in detail above, we have considered the locations of those 15 of the least deleterious variants that have known single-residue mutations, retain at least 10% residual activity and are not associated with CNSHA. With only a model of the human enzyme, we should not seek explanations for the ways in which the structure may alter to accommodate each of these changes. Nonetheless we have noted that there is a strong tendency for these variant residues to be accessible, on external helices or loops. Only one variant, 285Arg → His, involves mutation of an inaccessible residue that is well conserved between species, and our model will accommodate a histidine. In contrast, of the class I variant residues, 12 have very low accessibility and 2 of these and 5 others are totally conserved.

Among the less deleterious variants, G6PD A-, the common African deficient variant contains two mutations, 68Val → Met and 126Asn → Asp relative to G6PD B.\(^36\) The 126Asn → Asp mutation is found on its own in the nondeficient variant G6PD A. By contrast, the mutation 68Val → Met, which is the commonest in G6PD A-, has not yet been encountered on its own and when artificially engineered into normal G6PD B it does not cause deficiency.\(^38\) It has been concluded that the two mutations 126Asn → Asp and 68Val → Met act synergistically to cause enzyme deficiency. This so far unique finding is now justified by our three-dimensional model: indeed, we find that these two residues, on \(βB\) and \(α\), are spatially very close (some 8 Å apart, as illustrated in Fig 6). In the absence of the experimentally determined structure of the human enzyme, it is not reasonable to speculate further on the detailed differences in the sidechain interactions of the residues in the A- variant. However, we may surmise that the two substitutions interact specifically with each other causing disruption in the protein structure, possibly in the \(β\)-sheet of the coenzyme domain.

To understand in finer detail what amino acid changes may affect the dimer-tetramer equilibrium, and what weight must be assigned to different changes that may contribute to enzyme instability in the many known variants with differing severity of G6PD deficiency, it will be necessary to solve the three-dimensional structure of human G6PD. However, the structure modeled here on the basis of the \(L\) *mesenteroides* G6PD has already shown that residues in the active site
have not been modified in known variants, that a weakened dimer interaction is responsible for a large class of variants which give rise to CNSHA, and that residues distant in the primary structure and from different subunits are involved in the same sensitive region of the dimer interface.

ACKNOWLEDGMENT

We are grateful to Prof L.N. Johnson for facilities and support and to Robert Esnouf for use of his program BOBSCRIPT.

REFERENCES


34. Hirono A, Nakayama S, Fuji H, Miwa S: Molecular abnormality of a Japanese glucose-6-phosphate-dehydrogenase variant.
(G6PD Kobe) associated with a greatly increased affinity for galactose-6-phosphate. Am J Haematol 45:185, 1994


39. Town M, Bautista JM, Mason PJ, Luzzatto L: Both mutations in G6PD A− are necessary to produce the G6PD deficient phenotype. Hum Mol Genet 1:171, 1992


Glucose 6-phosphate dehydrogenase mutations causing enzyme deficiency in a model of the tertiary structure of the human enzyme

CE Naylor, P Rowland, AK Basak, S Gover, PJ Mason, JM Bautista, TJ Vulliamy, L Luzzatto and MJ Adams