Where is the weak linkage in the globin chain?

Viroj Wiwanitkit

Department of Laboratory Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

Abstract: Hemoglobinopathies are important inherited disorders with high prevalence in many tropical countries. Prediction of protein nanostructure and function is a great challenge in proteomics and structural genomics. Identifying the point vulnerable to mutation is a new trend in research on disorders at the genomic and proteomic level. A bioinformatics analysis was performed to determine the positions that tend to correspond with peptide motifs in the amino acid sequence of alpha and beta globin chains. To identify the weak linkage in alpha globin and beta globin chains, a new bioinformatics tool, GlobPlot, was used. For the alpha globin chain, 22 positions were identified: the disorders were found at positions 3–8, 38–42, 46–51, and 75–79. For the beta globin chain, 46 positions were identified: the disorders were found at positions 61–146. The study showed that weak linkages in alpha globin and beta globin chains can be identified and provide good information for predicting possible new mutations that could lead to new hemoglobinopathies.

Keywords: globin, hemoglobinopathy, protein structure, weak linkage

Introduction

Hemoglobinopathies are important inherited disorders with high prevalence in many tropical countries (Atweh et al 2003). In structurally abnormal hemoglobins, the following mechanisms can be invoked: single nucleotide base substitutions leading to amino acid replacement or chain termination variants; nucleotide deletions (or additions) leading to deletion and frameshift variants; and nonhomologous crossing over leading to the production of fused globin chains (Forget 1979). A defect in the globin chain is the main cause of hemoglobinopathies. The molecular basis of the hemoglobinopathies, disorders characterized by the absence of, or decreased, synthesis of alpha globin or beta globin chains, is quite heterogeneous (Forget 1979). The mutation in the gene coding for alpha globin or beta globin chains is the underlying pathogenesis of hemoglobinopathies (Atweh et al 2003). Clinical manifestations of these disorders vary from mild to severe presentations (Atweh et al 2003).

Predicting protein nanostructure and function is a great challenge in proteomics and structural genomics. Identifying the point vulnerable to mutation is a new trend in research on disorders at the genomic and proteomic level (Levin et al 2002; Lee and Wang 2005). Generally, disordered regions in proteins often contain short linear peptide motifs that are important for protein function. Identification of the peptide motifs in the amino acid sequence can give a good prediction of the weak linkages in a protein (Levin et al 2002; Lee and Wang 2005). A bioinformatics analysis was peformed to determine positions that tend to correspond with peptide motifs in the amino acid sequence of alpha globin and beta globin chains.

Materials and methods
Obtaining the amino acid sequence
The database ExPASy (Gasteiger et al 2003) was used for searching the amino acid sequence of normal human alpha globin and beta globin chains. The derived sequences were then used to identify the weak linkage.

Correspondence: Viroj Wiwanitkit
Department of Laboratory Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand 10330
Tel + 66 2 256 4163
Email: wviroj@yahoo.com

Identification of weak linkage in the globin chains

To identify the weak linkage in alpha globin and beta globin chains, a new bioinformatics tool, GlobPlot (Linding et al 2003), was used. GlobPlot is a web service that allows the user to plot the tendency within the query protein for order/globularity and disorder (Linding et al 2003). It successfully identifies interdomain segments containing linear motifs, and also apparently ordered regions that do not contain any recognized domains (Linding et al 2003).

Results

Twenty-two positions were identified for the alpha globin chain (Figure 1). The disorders were found at positions 3–8, 38–42, 46–51, and 75-79. Forty-six positions were identified for the beta globin chain (Figure 2). The disorders were found at positions 61–146.

mvLSPADKtn vkaawgkvga hageygaeal ekmflsfPTT KTyfpHFDLS Hgsaqvkghg kkvadaltna vahvDDMPNa lsalsdlhah klrvdpvnfk llshcllvtl aahlpaeftp avhasldkfl asvstvltsk yr

Figure 1 Identified positions (in capitals) that correspond with peptide motifs in amino acid sequence of alpha globin.

vhltpeeksa vtalwgkvnv devggealgr llvvypwtqr ffesfgdlst pdavmgnpkv KAHGKKVLGA FSDGLAHLDN LKGTFATLSE LHCDKLHVDP ENFRLLGNVL VCVLAHHFGK EFTPPVQAAY QKVVAGVANA LAHKYH

Figure 2 Identified positions (in capitals) that correspond with peptide motifs in amino acid sequence of beta globin.

Discussion

Structural aberration is believed to be the main underlying pathogenesis in hemoglobinopathy (Atweh et al 2003). Some disorders are described as a single substitution with other effects on the sequence frame; other disorders are mentioned as a frameshift (Atweh et al 2003). The mutations in globins can bring clinical manifestation in affected subjects. There are many reported mutations of globin and there are many undetermined ones. In the present study, an algorithm was used to identify the positions in the amino acid sequences of alpha globin and beta globin chains that can mutate.

Many positions were identified; some are known and others are newly discovered. Of interest, fewer positions were identified in alpha globin than in beta globin. More points with known mutations have been reported for beta globin than for alpha globin (Atweh et al 2003). The present study shows that weak linkages in the alpha globin and beta globin chains can be identified and can provide good information for predicting possible new mutations that can lead to new hemoglobin disorders.

References

Atweh GF, DeSimone J, Saunthararajah Y, et al. 2003. Hemoglobinopathies. *Hematology (Am Soc Hematol Educ Program)*, 1:14–39.

Forget BG. 1979. Molecular genetics of human hemoglobin synthesis. *Ann Intern Med*, 91:605–16.

Gasteiger E, Gattiker A, Hoogland C, et al. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*, 31:3784–8.

Lee C, Wang Q. 2005. Bioinformatics analysis of alternative splicing. *Brief Bioinform*, 6:23–33.

Levin JM, Penland RC, Stamps AT, et al. 2002. Using in silico biology to facilitate drug development. *Novartis Found Symp*, 247:222–3

Linding R, Russell RB, Neduva V, et al. 2003. GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res*, 31: 3701–8.