Meta-Analysis and Systematic Review

Journal of INTERNATIONAL MEDICAL RESEARCH

Journal of International Medical Research
48(6) 1–6

DOI: 10.1177/0300060520929853
journals.sagepub.com/home/imr
SAGE

High glycine content in TDP-43: a potential culprit in limbic-predominant age-related TDP-43 encephalopathy

Shanshan An[1,*], Xiaoxiao Zhang[2,*], Yunfan Shi[1,*], Jiaming Zhang[3], Yulin Wan[1], Yuchuan Wang[4], Ying Zhang[5] and Qiuyun Liu[1]

Abstract

Numerous risk factors for heart disease or dementia harbor over 10% valine plus glycine content. Interestingly, TDP-43 contains 6.0% valine and 13.3% glycine, and the buildup of this protein in the brains of patients with limbic-predominant age-related TDP-43 encephalopathy has dire consequences. The two γ-methyl groups in valine enable hyperconjugation, which enhances the van der Waals interaction between its side group and the carbonyl carbon. This extends the C=O bond length, and this weakened C=O bond augments the secondary chemical bonding of the carbonyl oxygen atom to cations. This, in turn, promotes the formation and buildup of insoluble and rigid salts such as calcium oxalate, which is postulated to be a major cause of heart disease. Similarly, the long C=O bond length in glycine results in a weakened C=O bond with an enhanced affinity toward cations and the formation of insoluble salts. Further, several prion proteins possess a high glycine content of approximately 20%. The insoluble calcium salts produced may promote aggregate formation via secondary chemical bonding between calcium and glycine, as well as between

[1]State Key Laboratory of Biocontrol, Biomedical Center, School of Life Sciences, Sun Yat-sen University, Guangzhou, China
[2]Yunnan Key Laboratory of Stem Cell and Regenerative Medicine, Biomedical Engineering Research Center, Kunming Medical University, Kunming, China
[3]School of Chemistry, Sun Yat-sen University, Guangzhou, China
[4]Center for Synthetic Biology Engineering Research, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, China
[5]Guangzhou Center for Disease Control and Prevention, Guangzhou, China

*These authors contributed equally to this work.

Corresponding author:
Qiuyun Liu, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China.
Email: lsslqy@mail.sysu.edu.cn

calcium and valine. Chemical and biochemical insights will help us to better understand the etiology of disorders linked to protein aggregates.

Keywords
TDP-43, limbic-predominant age-related TDP-43 encephalopathy, glycine, valine, cations, oxalate, calcium, insoluble, aggregate

Date received: 19 July 2019; accepted: 6 May 2020

Introduction

The repeated failures in drug development for Alzheimer's disease and dementia suggest that additional toolkits are needed in addition to the interrogation of biochemical pathways and disease pathology. Limbic-predominant age-related TDP-43 encephalopathy (LATE) is a common disease of the "oldest-old."[1] Many of the aggregated proteins that are causal factors of heart disease or Alzheimer's disease possess over 10% valine plus glycine (V + G) content. Methyl groups can be either electron-donating or electron-withdrawing, depending on the local milieu. The two γ-methyl groups in valine enable σ-σ hyperconjugation and electron delocalization. The increased electron density in C_β-H enhances potent van der Waals interactions between the side group and carbonyl carbon.[2] This strengthened secondary bonding with the carbonyl carbon results in extension of the C=O bond length,[3] and this weakened C=O bond then augments the secondary chemical bonding of the carbonyl oxygen atom to cations, particularly divalent cations.[2] This attribute of valine residues enables its affinity toward calcium and enhances the formation of insoluble and rigid calcium salts such as calcium oxalate, which has been postulated as a major cause of heart disease.[2] Because ethanol and acetic acid are structurally similar to oxalate, red wines reduce the risk of heart disease and dementia and can extend one's lifespan, possibly via the inhibition of oxalate generation (Figure 1).[4] In older people who have attenuated respiratory chain activities, the Krebs cycle and its shunt produce excess protons and organic acids such as oxalate (Figure 1). Alzheimer's disease and cancer are mutually protective. This can be explained by calcium supplements substantially reducing cancer risk, whereas local strong acids help dissolve insoluble and rigid calcium oxalate and other insoluble salts.[5] The C=O bond length in glycine is slightly longer than its counterpart in valine,[3] and this weakened C=O bond enhances the formation of insoluble salts, triggering aggregation. Some prion proteins possess around 20% glycine content. Interestingly, TDP-43 contains 6.0% valine and 13.3% glycine, and the buildup of this protein in patients with LATE has negative effects on the brain. The primary structure of TDP-43 is bipartite. The C-terminal fragment of TDP-43 is particularly glycine-rich at 27.0% while being poor in basic amino acids at 2.8% (Table 1). The N-terminal fragment of TDP-43 possesses 15.0% basic amino acids and 15.4% V+G residues (Table 1). The higher-order structure of TDP-43 can bring positively charged basic amino acids, glycine, and valine together (Figures 2 and 3),[6–8] enhancing the formation of calcium oxalate and, in turn, the formation of

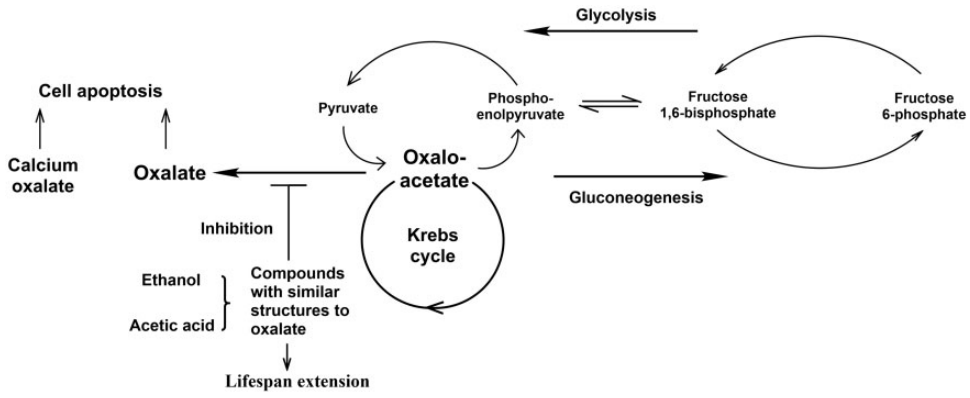

Figure 1. Biochemical pathway leading to the generation of oxaloacetate and oxalate.

Table 1. Amino acid content of the TDP-43 protein.

Position of protein residues	Number of glycine (G) residues	Number of valine (V) residues	V + G content (%)	Number of arginine (R) residues	Number of lysine (K) residues	Number of histidine (H) residues	R + K + H content (%)
1–273	17	25	15.4	17	19	5	15.0
274–414	38	0	27.0	3	1	0	2.8
1–414	55	25	19.3	20	20	5	10.9

aggregates via secondary chemical bonding between calcium (in the form of insoluble salts) and glycine, and between calcium and valine. Other protein risk factors for LATE also possess relatively high V + G content, possibly contributing to the formation of insoluble and rigid salts such as calcium oxalate. Table 2 shows the glycine, valine, and basic amino acid content of the four causative factors of Alzheimer's disease, indicating that the V + G percentage is also high in these proteins. Amino acid polymorphisms of causative factors promoting the traffic of either divalent cations or negatively charged oxalate can have a significant impact on neurodegenerative diseases. For example, the ApoE4 polymorphism confers high risk for Alzheimer's disease and atherosclerosis, with homozygous carriers (arg/arg) having a higher risk than heterozygous carriers.[9] The positively charged arginine residues in the ApoE4 polymorphic site enhance the traffic of anions such as oxalate, consequently augmenting the formation of calcium oxalate.

A previous study showed that the N-terminal domain of prion protein, PrP^{C}, contains repeats of PHGGGWGQ in five to six sites that bind divalent Cu^{2+} via glycine chelation.[10] High glycine content has also been reported for the causative factors of amyotrophic lateral sclerosis, which include SOD1, TARDBP, and FUS, ranging from around 10% to 30%. Oxaloacetate is metabolized to phosphoenolpyruvate (PEP) by PEP carboxykinase (PEPCK), and the overexpression of PEPCK can confer an extended lifespan,[11] perhaps by channeling oxaloacetate toward PEP formation and thus minimizing oxalate buildup. Further, calcium oxalate is a primary component of renal stones.

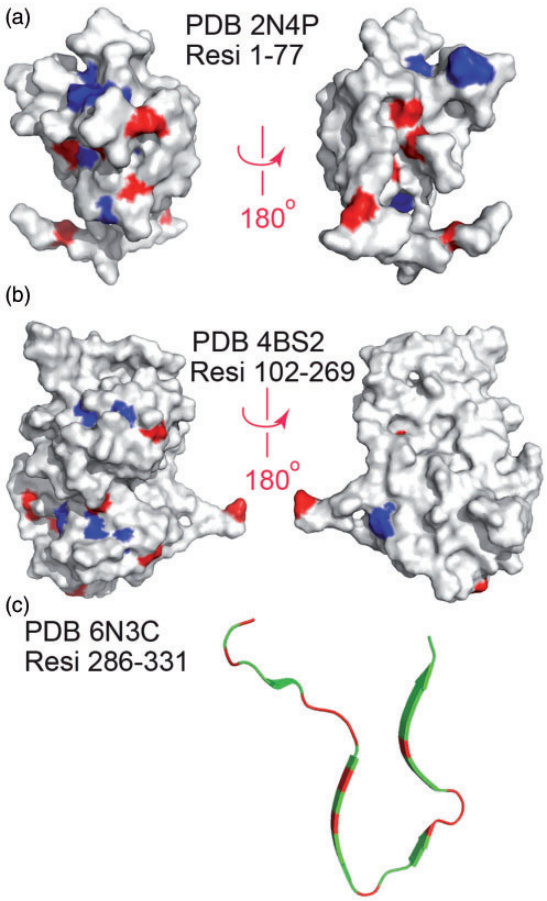

Figure 2. The distribution of valine and glycine residues on the TDP-43 protein. No complete protein structure is available for TDP-43; therefore, the protein structure of its three regions are shown separately. (a) Residues 1 to 77, PDB 2N4P;[6] (b) residues 102 to 269, PDB 4BS2;[7] (c) residues 286 to 331, PDB 6N3C;[8] which together cover most of TDP-43 polypeptide. The valine and glycine residues are depicted in blue and red, respectively. In (a) and (b), the protein regions were rendered in surface representation, with the front and back faces shown. In (c), the structure is shown in the ribbon representation and only one view angle is provided.

These phenomena corroborate the hypothesis that the formation of insoluble salts may be destructive to cells.

A high proportion of one or more particular amino acids in proteins is not rare. It is a hallmark of many risk factors for human diseases and numerous virulence factors in viruses. For instance, red meat is thought to be mutagenic, and it is marked by the presence of myoglobin which possesses around 20% to 21% basic amino acids, attracting anions such as Cl^- and contributing to the local formation of HCl, which generates mutations.[12]

A carbohydrate/vitamin diet taken at intervals and under the guidance of a physician does not provide essential amino acids[13,14] and can be used to reduce a patient's level of full-length, disease-causative factors and their fragments, thereby

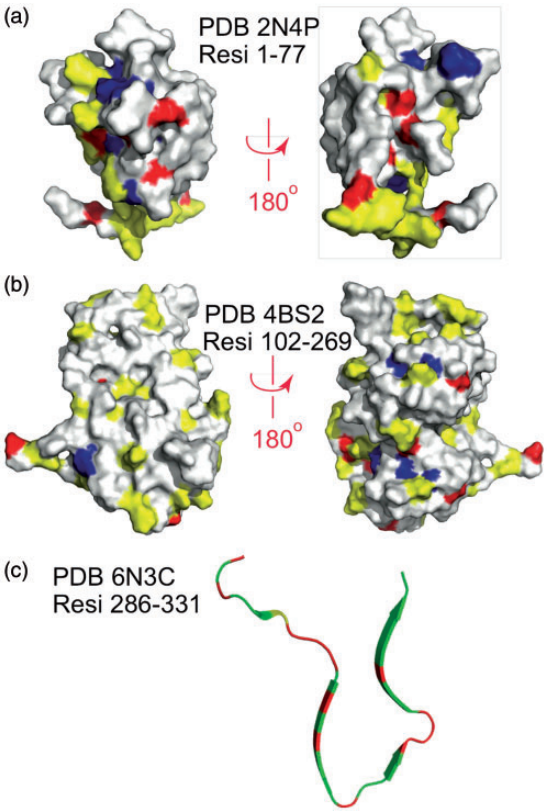

Figure 3. The distribution of the valine, glycine, and basic amino acid residues on the TDP-43 protein. The valine and glycine residues are depicted in blue and red, respectively, while the basic residues (histidine, arginine, and lysine) are represented in yellow.

Table 2. Amino acid content of the causative protein factors of Alzheimer's disease.

Gene	Number of amino acid residues	Number of V + G residues	V + G content (%)	Number of R + K + H residues	R + K + H content (%)
APP	770	103	13.4	103	13.4
S182	467	64	13.7	46	9.9
STM2	448	68	15.2	32	7.1
APOE4	317	42	13.2	50	15.8

addressing symptoms caused by the over-representation of essential and/or non-essential amino acids. A plant-based diet also reduces the intake of essential amino acids. Reduced food consumption or occasional fasting may also be beneficial to patients carrying protein risk factors with high proportions of non-essential amino acids such as glycine. However, this should only be conducted under the supervision of medical professionals. Caution should be exercised in the treatment of diseases

with food regimens, as these need to be carefully designed.

In summary, the secondary chemical bonding of the carbonyl oxygen in glycine residues is critical in the etiology of LATE, and this phenomenon may also contribute, at least in part, to the pathogenesis of other human disorders with glycine-rich causative factors, such as amyotrophic lateral sclerosis and prion diseases.

Acknowledgements

We thank Wei Xie, Tao Gan, Li Xu, and Weiguo Cao for discussions and Yan Shi for editing.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by grants from the Science and Technology Transformation Program of Sun Yat-sen University of China (33000-18843234), the Guangzhou Science and Technology Program (201804010328) to Q. Liu, and the National Science and Technology Major Project of the Ministry of Science and Technology of China (2017ZX10103011) to Y. Zhang.

ORCID iD

Qiuyun Liu https://orcid.org/0000-0001-5533-0128

References

1. Nelson PT, Dickson DW, Trojanowski JQ, et al. Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. *Brain* 2019; 142: 1503–1527.
2. Qi J, Zhang X, Chen Z, et al. Transgenic proteins rich in valine or glycine are concerns for heart disease patients. *Eur J Prev Cardiol* 2018; 25: 883–884.
3. Gu Z, Zambrano R and McDermott A. Hydrogen bonding of carboxyl groups in solid-state amino acids and peptides: comparison of carbon chemical shielding, infrared frequencies, and structures. *J Am Chem Soc* 1994; 116: 6368–6372.
4. An S, Li X, Tang M, et al. Functional duality of ethanol on cancer. *Med Hypotheses* 2019; 122: 124–125.
5. Wan Y, Hu X and Liu Q. Protection of cancer in patients with neurodegenerative diseases. *Eur J Cancer Prev* 2019; 28: 459.
6. Mompean M, Romano V, Pantoja-Uceda D, et al. The TDP-43 N-terminal domain structure at high resolution. *FEBS J* 2016; 283: 1242–1260.
7. Lukavsky PJ, Daujotyte D, Tollervey JR, et al. Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat Struct Mol Biol* 2013; 20: 1443–1449.
8. Cao Q, Boyer DR, Sawaya MR, et al. Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat Struct Mol Biol* 2019; 26: 619–627.
9. Iron A, Richard P, Beucler I, et al. Pathology of the human apolipoprotein E gene. *Bull Acad Natl Med* 1994; 178: 415–426.
10. Brown DR, Qin K, Herms JW, et al. The cellular prion protein binds copper in vivo. *Nature* 1997; 390: 684–687.
11. Hakimi P, Yang J, Casadesus G, et al. Overexpression of the cytosolic form of phosphoenolpyruvate carboxykinase (GTP) in skeletal muscle repatterns energy metabolism in the mouse. *J Biol Chem* 2007; 282: 32844–32855.
12. Tang M, Zhou Y, Li Y, et al. Hydrogen donors and acceptors and basic amino acids jointly contribute to carcinogenesis. *Med Hypotheses* 2017; 98: 42–44.
13. Gan T, Fu M, Wu J, et al. How to design carbohydrate diet regimens for heart disease patients. *Eur J Prev Cardiol* 2018; 25: 979–980.
14. He R, He S, Li X, et al. How can heart disease patients prevent complications from viral infections? *Eur J Prev Cardiol* 2018; 25: 758.