ISSN 0006-2979, Biochemistry (Moscow), 2010, Vol. 75, No. 1, pp. 110-114. © Pleiades Publishing, Ltd., 2010.
Published in Russian in Biokhimiya, 2010, Vol. 75, No. 1, pp. 136-141.

Identifying Functional Residues in *Arabidopsis thaliana* Zeta Class Glutathione *S*-Transferase through Screening Inactive Point Mutants

Xiwen Chen[#], Jia Liu[#], Peng Yang, and Defu Chen*

Laboratory of Molecular Genetics, College of Life Sciences, Nankai University, Tianjin 300071, P. R. China; fax: 86-22-2350-0133; E-mail: chendefu@nankai.edu.cn

Received March 11, 2009
Revision received April 20, 2009

Abstract—The functional residues of z-class glutathione *S*-transferase were identified by screening inactive point mutants from a random mutagenesis library. First, a random mutant library was constructed using error-prone polymerase chain reaction, and then candidate inactive mutants were screened by a high-throughput colorimetric assay. Twenty-five mutants were obtained, and 12 that formed inclusion bodies were discarded. The remaining 13 mutants that expressed soluble protein were used for accurate quantification of enzymatic activity and sequencing. The mutants W15R, C19Y, R22H/K83E, P61S, S73P, S109P, and Q112R were found to have activity lower than 1% of the wild-type and were considered as "inactive mutants", whereas the mutants K83E, Q102R, and L147F still have a large fraction of the activity and were thus considered as "partially inactivated mutants". Molecular modeling experiments disclosed that mutations resulting in inactivation of the enzyme were found in or near the binding pocket, whereas mutations resulting in partial inactivation were distant from both substrates. The role of the residue Ser73 in the enzyme was verified by site-directed mutagenesis. The result suggested that screening inactive point mutants from a random mutagenesis library is an efficient way of identifying functional residues in enzymes.

DOI: 10.1134/S0006297910010141

Key words: functional residues, random mutation, site-directed mutagenesis, *Arabidopsis thaliana*, z-class glutathione *S*-transferase

Identifying functional residues frequently requires a large amount of information input, including sequence alignments with homologous proteins, known three-dimensional structures, and mutational analysis for desired residues. More and more evidence has demonstrated that mutations at functionally important residues are more likely to inactivate enzymes [1-4]. Correspondingly, inactive mutations have large possibility to occur in functional residues. If an appropriate mutation frequency is controlled and a single amino acid substitution per protein is introduced, the inactive point mutations could, therefore, be used for identifying the functional residues in enzymes.

The glutathione *S*-transferases (GSTs) (EC 2.5.1.18) are ubiquitous enzymes, existing as multigene families in bacteria, fungi, plants, and animals [5]. They could catalyze the nucleophilic attack of the tripeptide glutathione (GSH) on the electrophilic center of hydrophobic compounds [6]. Based on gene organization and amino acid sequence similarity, cytosolic plant GSTs are categorized into four classes: φ, τ, θ, and z [5]. The newly developed z-class GST (GSTZ) can efficiently catalyze the GSH-dependent dechlorination of dichloroacetic acid (DCA), a potential carcinogenic contaminant in chlorinated drinking water [7], and could also improve the transgenic cotton recuperator line and enhance the yield of the hybrid generation [8].

However, GSTZ in *Arabidopsis thaliana* (*At*GSTZ) presented dramatically different properties from other GSTs. It demonstrated no activity toward the traditional substrates of GST, but showed unique activities toward DCA and maleyl acetone [9]. Furthermore, *S*-hexylglutathione, a GSH conjugate known to inhibit other GSH-dependent enzymes, had no significant effect on its activ-

Abbreviations: *At*GSTZ, *Arabidopsis thaliana* zeta class glutathione *S*-transferase; DCA, dichloroacetic acid; DCA-DC, DCA dechlorination; EP-PCR, error-prone polymerase chain reaction; GSH, glutathione; GST, glutathione *S*-transferase; IPTG, isopropyl β-D-thiogalactopyranoside.

[#] Contributed equally to this work.

* To whom correspondence should be addressed.

ity [9]. Although the active-site residue Ser17 was identified as a contributor to the stabilization of GSH [10], its contribution was not as profound as seen in other GSTs [11]. This suggested that other residues might also play important roles in catalysis. In this sense, it becomes an imperative work to probe functional residues in *At*GSTZ.

In this study, *At*GSTZ was subjected to random mutagenesis, and inactive point mutants were screened by a high-throughput assay. Several functional residues of *At*GSTZ were identified. The result suggested that screening inactive point mutants from a random mutagenesis library is an efficient way of identifying functional residues in enzymes, which would facilitate in-depth study in structure–function relationship of proteins and facilitate protein engineering.

MATERIALS AND METHODS

Plasmid and chemicals. The recombinant plasmid pET21b(+)-*AtGSTZ* was constructed in our previous study [12]. Isopropyl β-D-thiogalactopyranoside (IPTG), *Nde* I, *Not* I, and all PCR chemicals were purchased from TaKaRa Biotechnology Co. Ltd. (China). GSH was purchased from Sangon Co. Ltd. (China). Other chemicals were purchased from common commercial suppliers.

Construction of *AtGSTZ* mutant library. Random mutagenesis of *AtGSTZ* was performed using EP-PCR (error-prone polymerase chain reaction) [13]. To obtain an appropriate mutation frequency, the conditions of EP-PCR were optimized as follows: 25 μl reaction mixture containing 0.25 ng recombinant plasmid pET21b(+)-*AtGSTZ*, 0.5 U *rTaq* polymerase, 0.2 mM dNTPs, 2.5 mM $MnCl_2$, 0.12 μM of primer pET-T7 (5′-GAAATTAATACGACTCACTATAGGGGAA-3′) and primer Seq-Rew (5′-CGGATATAGTTCCTCCTTTCAGC-3′) each; 94°C for 5 min, 40 cycles of 94°C for 40 sec, 58°C for 32 sec, and 72°C for 90 sec, and finally 72°C for 7 min. The reaction was performed by a TaKaRa PCR Thermal Cycler Dice TP600. PCR products were purified by gel electrophoresis and then digested with *Nde* I/*Not* I. The resulting fragments were ligated with pET21b(+) that was treated with the same restriction enzymes. The recombinant plasmids containing mutated *AtGSTZ* gene were then transformed into *Escherichia coli* DH5α-FT. Their plasmids were extracted and introduced into *E. coli* BL21(DE3).

Screening for inactive mutants. On the basis of the method described by Tong et al. [14], a high-throughput colorimetric method for screening candidate inactive mutants was developed as follows. Each single colony of BL21(DE3) transformant was inoculated into 150 μl of 2× TY liquid medium containing 50 μg/ml ampicillin and 1 mM IPTG. After overnight incubation with shaking, it was frozen at −70°C and thawed at 4°C for three cycles. Then 30 μl of 19 mM NaH_2PO_4/81 mM Na_2HPO_4 buffer (pH 7.4) containing 3.3 mM DCA and 13.3 mM GSH was added. The mixture was incubated at 37°C for 40 min, and then 65 μl of 408 mM NaH_2PO_4/392 mM Na_2HPO_4 buffer (pH 6.8) and 85 μl of fresh phenylhydrazine hydrochloride solution (100 mg in 15 ml of water) were added. The mixture was kept at room temperature for 10 min and then placed on ice for 10 min. Thereafter, 210 μl of chilled concentrated HCl, and 85 μl of fresh potassium ferricyanide solution (500 mg in 15 ml of water) were added to each mixture. Colonies, which showed no visible color changes, were selected. PCR and restriction analysis was performed to exclude ligation with no insert. The verified clones were sequenced by Beijing Sunbiotech Inc. (China). The sequenced mutants were induced with 1 mM IPTG at A_{600} = 0.8 and expressed at 37°C overnight. A 12.5% SDS-PAGE analysis was performed to exclude insoluble mutants. Enzyme expression, purification, and quantitative activity assay were performed on soluble mutants as previously described [12].

Molecular modeling of *At*GSTZ. The substrate-superimposed model of *At*GSTZ was performed using ICM-Pro 3.4-8 (MolSoft L. L. C., USA) [15, 16]. The GSH and sulfate ion (SO_4^{2-}) from human z-class GST (*h*GSTZ, PDB entry 1FW1) [17] were superimposed into the apo-form crystal structure of *At*GSTZ monomer (PDB entry 1E6B) [10].

Site-directed mutagenesis of Ser73 in *At*GSTZ. Site-directed mutagenesis of Ser73 was performed by overlap extension PCR [18]. The upstream fragment was amplified using primer pET-T7 and reverse mutagenic primer (5′-GCAAAcgcATCATTAATCACAACATCTCCATC-3′, mutated codon lowercased). The downstream fragment was amplified using primer Seq-Rew and forward mutagenic primer (5′-AATGATgcgTTTGCGATAATAATGTATCTGGA-3′, mutated codon lowercased). The amplifications were performed in 15 μl reaction mixture containing 0.15 ng pET21b(+)-*AtGSTZ*, 1 U of *PrimerStar* DNA polymerase, 0.3 μM of each primer, 0.2 mM of each dNTP under the above cycling conditions. The PCR products of the two fragments with slight overlap were purified and used as template in overlap extension, which was performed using *PrimerStar* DNA polymerase and primers pET-T7 and Seq-Rew under the following conditions: 94°C for 5 min, 5 cycles of 94°C for 60 sec, 54°C for 60 sec, and 72°C for 60 sec, followed by 30 cycles of 94°C for 40 sec, 58°C for 32 sec, and 72°C for 90 sec, and a final extension of 72°C for 7 min. The PCR products of the entire gene with mutations at Ser73 were purified, digested, ligated, and transformed as above and confirmed by sequencing.

RESULTS

Screening of candidate inactive mutants by activity. The mutation frequency of EP-PCR varies according to

different amount of initiating template, concentration of $MnCl_2$, and cycling numbers. These parameters were optimized first as described in "Materials and Methods" to introduce a single amino acid substitution per enzyme. Sequencing of the selected mutants identified 31 point mutations in 25 mutants with the average amino acid substitution per enzyme being 1.24, which broadly matched the expected single mutation.

A total of 3000 colonies were screened by a high-throughput colorimetric activity assay. Thirty-two clones that showed no visible changes in color were obtained. Plasmid PCR and restriction analysis found seven of them contained no insert. The remaining 25 mutants were kept for sequencing and further quantification of DCA-DC (DCA dechlorination) activity.

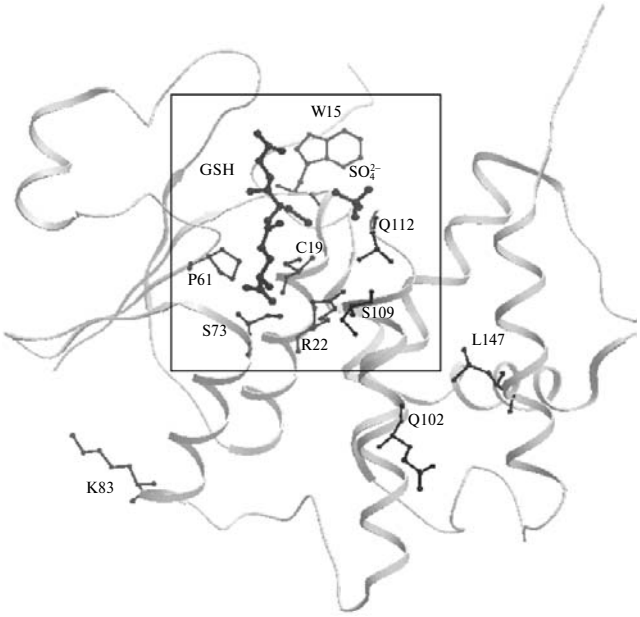
The DCA-DC activity of the 25 mutants is shown in Table 1. Among them, 12 mutants expressed mainly as inclusion bodies and no activity was detected. Soluble mutants W15R, C19Y, R22H/K83E, P61S, S73P, S109P, and Q112R were found to have less than 1% of the wild-type activity and were defined as "inactive mutants". Soluble mutants K83E, Q102R, and L147F retained 49.4, 31.6, and 52.3% of wild-type activity, respectively, and were defined as "partially inactive mutants". The possible reason that the accurate activities of K83E, Q102R, and L147F were different from those determined in high-throughput activity assay might lie in their abnormal expression in high-throughput activity assay.

Comparing the activity of some mutants could provide useful hints in the functional residues of the enzyme. The mutant L147P resulted in inclusion bodies, whereas L147F retained 52.3% activity, suggesting that a hydrophobic residue at position 147 might be necessary for DCA-DC activity. The mutant K83E showed 49.4% activity, while the mutant K83E/R22H demonstrated <1% activity, suggesting that Arg22 might be catalytically important, which is in accordance with the notion that Arg22 participates in formation of a positively charged binding pocket and contributes to substrate binding [10].

Table 1. Effects of different mutations on dichloroacetic acid dechlorination (DCA-DC) activity

Mutants*	DCA-DC activity**, %
L11P (2), L11P/I140T, L27P (3), K30E, L80P/N116S, L147P, L150P, L169R, S105R/Q213R	inclusion bodies, no detectable activity
W15R (2), C19Y, C19R/S105G, R22H/K83E (2), P61S, S73P, S109P, Q112R	< 1.0
K83E	49.4 ± 0.1
Q102R	31.6 ± 0.4
L147F	52.3 ± 0.9

* Numbers in parentheses reflect frequency of identical mutants.
** Relative to wild-type enzyme activity. Data are shown as means ± SD (n = 3).



Distribution of the residues whose mutations resulted in inactivation. Residues in or near the binding pocket are displayed in the box. The backbone of *At*GSTZ is displayed in ribbon. Residues and substrate molecules are presented in ball and stick

Distribution of residues resulting in enzyme inactivation. To determine the distribution of the residues whose mutation resulted in inactivation in the enzyme, the substrates from *h*GSTZ were superimposed into the apoform *At*GSTZ. *h*GSTZ was chosen as a reference because it belongs to z-class GST and exhibits a sequence identity of 48% with *At*GSTZ. As shown in the figure, Trp15, Cys19, Arg22, Pro61, Ser73, Ser109, and Gln112, the mutations of which resulted in inactivation of the enzyme, were found in or near the binding pocket, whereas Lys83, Gln102, and Leu147, the mutations of which resulted in partial inactivation, were distant from both substrates.

The distances between the substrates and the respective residues in these mutated enzymes were also calculated. Because the accurate orientation of each side chain is not known for the *in silico* model, the distance between the substrates and the nearest atoms instead of the potential reactive atoms were calculated as shown in Table 2. Trp15, Ser109, and Gln112, the mutations of which resulted in partial inactivation, were found to be far (8.3, 7.1, and 6.5 Å, respectively) from GSH, whereas Cys19, Arg22, Pro61, and Ser73, the mutations of which resulted in inactivation of the enzyme, were found to be close (3.1, 3.1, 4.0, and 2.1 Å, respectively) to the substrate

Table 2. Distances between the amino acid residues and superimposed substrates

Residue	Distance*, Å	
	GSH	SO_4^{2-} (mimicking DCA)
W15	8.3 (C and Cys-Cβ)	7.1 (O and O^1)
C19	3.1 (N and Cys-Sγ)	4.8 (N and O^3)
R22	3.1 ($N\eta^1$ and γGlu-O^1)	6.0 ($N\eta^2$ and O^3)
P61	4 (Cδ and γGlu-O^2)	8.6 (Cδ and O^4)
S73	2.1 (Oγ and γGlu-O^1)	9.3 (Oγ and O^3)
S109	7.1 (N and γGlu-N)	11 (N and O^3)
Q112	6.5 (Cβ and γGlu-Cβ)	6.0 (O and O^3)

* Respective nearest enzyme and substrate atoms are indicated in parentheses.

GSH. However, no correlation between the distance of the mutated residues from DCA and the activity of the mutated enzyme was found. As the Oγ in the hydroxyl group of Ser73 was only 2.1 Å from the substrate GSH, it might be important in binding and stabilizing GSH.

Verification of the role of Ser73. To verify the role of Ser73, site-directed mutagenesis was performed to substitute Ser73 with alanine. The mutant S73A showed 21.6% of wild-type activity, suggesting that the hydroxyl group at 73 site cannot be essential but it catalytically important, which was consistent with the role of equivalent residues in other GSTs [19-23]. In addition, the mutant S73P showed a greater loss in activity, to only 0.1% of the wild-type activity when compared to S73A. This extra loss might be caused either by the misfolding of native structure or by the steric hindrance introduced by the bulky side chain of proline. However, S73P expressed as soluble form indicating that the steric hindrance mainly caused the inactivation of S73P.

DISCUSSION

Site-directed mutagenesis has been widely used to identify functional residues in enzymes [24]. Recent strategy also included an alanine scanning mutagenesis, which performed a set of site-directed mutagenesis on targeted fragments to determine substrate-binding region [25]. However, these strategies usually required considerable input of biochemical and biophysical information to target mutagenesis to a single residue or fragment in the full-length gene. Enzymatic theory indicated that the enzyme function is normally determined by a few special residues that constitute the active site. When these residues are mutated, the enzyme structure and function will change accordingly. If an appropriate mutation frequency is controlled and a single amino acid substitution per protein is introduced, inactive mutations are largely possible to occur in the functional residues. Thus, screening the inactive point mutants could reveal the functional residues of enzymes. Once the structural information is obtained, the enzyme could be engineered through target mutagenesis. This would enormously reduce the screening load and facilitate the engineering of desired enzymes. Thus, identifying functional residues is the first step for efficiently engineering enzymes.

Based on this assumption, *At*GSTZ was subjected to random mutagenesis, and inactive point mutants were screened by a high-throughput assay to identify functional residues in this study. A random mutant library was first constructed using EP-PCR in the presence of $MnCl_2$. EP-PCR was chosen because it is a widely used method in random mutagenesis, which has the lowest ability to create mutational diversity on the gene level [26]. The added Mn^{2+} ion can interact with the phosphate and/or base sites of DNA by displacing a proton to configure the structure of base–M–base [1, 2]. This is quite important because multiple mutations can introduce synergetic effect [27] and obscure the role of single residues.

Then a high-throughput colorimetric assay was developed for screening inactive mutants based on the dechlorinating activity against DCA. DCA dechlorinating activity was chosen because DCA is a more stable substrate than the natural substrate of maleylacetoacetate isomerase, and the reaction mixture turns pink when catalysis occurs. Thus the assay is much more straightforward and it is easier to develop a high-throughput assay method. After the candidate inactive mutants were obtained, their enzymes were extracted, purified and their accurate activities were determined. Two kinds of mutants were obtained. Among them, seven (W15R, C19Y, R22H/K83E, P61S, S73P, S109P, and Q112R) have less than 1% of the wild-type activity, whereas three (K83E, Q102R, and L147F) retained 31.6-52.3% of wild-type activity. Further analysis revealed that the position of mutation significantly affects the activity of the mutated enzyme. The mutations occurred in or near the binding pocket of the enzyme resulted in inactivation, whereas the mutations distant from them only partially inactivated their activity. This is in accordance with the notion that binding-pocket mutations tend to affect enzyme activity more than distant mutations [28]. Trp15, Cys19, Arg22, Pro61, Ser73, and Gln112 were all functional residues detected in the crystal structure of *At*GSTZ [10]. Trp15 was proposed to position Pro216 and thus direct the C-terminus away from the active site. Cys19 was demonstrated to influence catalysis but was not essential. Arg22 could form a positively charged locus in the active site, which might contribute to substrate binding. Pro61 and Ser73 were highly conserved residues in all GSTs and essential for catalysis. Gln112 appeared in the conserved motif IQP (residues 111-113), which is required for substrate binding or catalytic activity. This significant obser-

vation demonstrated that inactive point mutations could serve as indicators for functional residues. However, no inactive point mutant was detected in the active-site residue Ser17 in this study. This could be because mutagenesis occurred randomly, and it is not in conflict with our conclusion.

Identification of functional residues could be simply achieved by screening the inactive point mutants from a random mutagenesis library with a fine-tuned mutation frequency. As inactive mutants usually have distinct phenotype, it is much easier to screen them from the mutant library by a high-throughput activity assay. Therefore, compared to site-directed mutagenesis, inactive point-mutation screening demands the minimal information input. It dramatically facilitates in-depth study of functional residues in proteins/enzymes and thus opens the way to protein engineering.

The authors are grateful to Mr. Xin Dong and Ms. Su-dan Tao at Nankai University for their skillful help on screening mutants.

This work was supported by the National Natural Science Foundation of China (No. 30671183).

REFERENCES

1. Hirano, S., Mihara, K., Yamazaki, Y., Kamikubo, H., Imamoto, Y., and Kataoka, M. (2002) *Proteins*, **49**, 255-265.
2. Zhang, H., Huang, S., Feng, Y., Guo, P., and Jing, G. (2005) *Arch. Biochem. Biophys.*, **441**, 123-131.
3. Fukuma, K., Neuls, E. D., Ryberg, J. M., Suh, D. Y., and Sankawa, U. (2007) *J. Biochem.*, **142**, 731-739.
4. Brown, B. M., and Sauer, R. T. (1999) *Proc. Natl. Acad. Sci. USA*, **96**, 1983-1988.
5. Dixon, D. P., Lapthorn, A., and Edwards, R. (2002) *Genome Biol.*, **3**, 1-10.
6. Armstrong, R. N. (1997) *Chem. Res. Toxicol.*, **10**, 2-18.
7. Weisel, C. P., Kim, H., Haltmeier, P., and Klotz, J. B. (1998) *Environ. Health Perspect.*, **107**, 103-110.
8. Wang, X. D., and Li, Y. Y. (2002) *Sci. Agric. Sin.*, **35**, 137-141.
9. Dixon, D. P., Cole, D. J., and Edwards, R. (2000) *Arch. Biochem. Biophys.*, **384**, 407-412.
10. Thom, R., Dixon, D. P., Edwards, R., Cole, D. J., and Lapthorn, A. J. (2001) *J. Mol. Biol.*, **308**, 949-962.
11. Caccuri, A. M., Antonini, G., Nicotra, M., Battistoni, A., Lo Bello, M., Board, P. G., Parker, M. W., and Ricci, G. (1997) *J. Biol. Chem.*, **272**, 29681-29686.
12. Chen, D., Kawarasaki, Y., Nakano, H., and Yamane, T. (2003) *J. Biosci. Bioeng.*, **95**, 594-600.
13. Cirino, P. C., Mayer, K. M., and Umeno, D. (2003) *Meth. Mol. Biol.*, **231**, 3-9.
14. Tong, Z., Board, P. G., and Anders, M. W. (1998) *Biochem. J.*, **331**, 371-374.
15. Abagyan, R., Totrov, M., and Kuznetsov, D. N. (1994) *J. Comput. Chem.*, **15**, 488-506.
16. Abagyan, R., and Totrov, M. (1994) *J. Mol. Biol.*, **235**, 983-1002.
17. Polekhina, G., Board, P. G., Blackburn, A. C., and Parker, M. W. (2001) *Biochemistry*, **40**, 1567-1576.
18. Nagy, Z. B., Varga-Orvos, Z., Szakal, B., Tamas, L., and Puskas, L. G. (2006) *Anal. Biochem.*, **351**, 311-313.
19. Zeng, Q. Y., and Wang, X. R. (2005) *FEBS Lett.*, **579**, 2657-2662.
20. Winayanuwattikun, P., and Ketterman, A. J. (2004) *Biochem. J.*, **382**, 751-757.
21. Sun, Y. J., Kuan, I. C., Tam, M. F., and Hsiao, C. D. (1998) *J. Mol. Biol.*, **278**, 239-252.
22. Cardoso, R. M. F., Daniels, D. S., Bruns, C. M., and Tainer, J. A. (2003) *Proteins*, **51**, 137-146.
23. Fritz-Wolf, K., Becker, A., Rahlfs, S., Harwaldt, P., Schirmer, R. H., Kabsch, W., and Becker, K. (2003) *Proc. Natl. Acad. Sci. USA*, **100**, 13821-13826.
24. Liu, L., and Lomonossoff, G. P. (2008) *Meth. Mol. Biol.*, **451**, 395-404.
25. Huang, Y., Smith, C. A., Song, H., Morgan, B. P., Abagyan, R., and Tomlinson, S. (2005) *J. Biol. Chem.*, **280**, 34073-34079.
26. Zhao, H., Moore, J. C., Volkov, A. A., and Arnold, F. H. (1999) in *Manual of Industrial Microbiology and Biotechnology*, ASM, Washington, DC, pp. 597-604.
27. Bloom, J. D., Meyer, M. M., Meinhold, P., Otey, C. R., MacMillan, D., and Arnold, F. H. (2005) *Curr. Opin. Struct. Biol.*, **15**, 447-452.
28. Park, S., Morley, K. L., Horsman, G. P., Holmquist, M., Hult, K., and Kazlauskas, R. J. (2005) *Chem. Biol.*, **12**, 45-54.