Modern diversification of the amino acid repertoire driven by oxygen

Matthias Granold[a], Parvana Hajieva[b], Monica Ioana Toşa[c], Florin-Dan Irimie[c], and Bernd Moosmann[a,1]

[a]Evolutionary Biochemistry and Redox Medicine, Institute for Pathobiochemistry, University Medical Center of the Johannes Gutenberg University, 55128 Mainz, Germany; [b]Cellular Adaptation Group, Institute for Pathobiochemistry, University Medical Center of the Johannes Gutenberg University, 55128 Mainz, Germany; and [c]Group of Biocatalysis and Biotransformations, Faculty of Chemistry and Chemical Engineering, Babeş-Bolyai University, Cluj-Napoca 400028, Romania

Edited by Harry B. Gray, California Institute of Technology, Pasadena, CA, and approved November 21, 2017 (received for review October 1, 2017)

All extant life employs the same 20 amino acids for protein biosynthesis. Studies on the number of amino acids necessary to produce a foldable and catalytically active polypeptide have shown that a basis set of 7–13 amino acids is sufficient to build major structural elements of modern proteins. Hence, the reasons for the evolutionary selection of the current 20 amino acids out of a much larger available pool have remained elusive. Here, we have analyzed the quantum chemistry of all proteinogenic and various prebiotic amino acids. We find that the energetic HOMO–LUMO gap, a correlate of chemical reactivity, becomes incrementally closer in modern amino acids, reaching the level of specialized redox cofactors in the late amino acids tryptophan and selenocysteine. We show that the arising prediction of a higher reactivity of the more recently added amino acids is correct as regards various free radicals, particularly oxygen-derived peroxyl radicals. Moreover, we demonstrate an immediate survival benefit conferred by the enhanced redox reactivity of the modern amino acids tyrosine and tryptophan in oxidatively stressed cells. Our data indicate that in demanding building blocks with more versatile redox chemistry, biospheric molecular oxygen triggered the selective fixation of the last amino acids in the genetic code. Thus, functional rather than structural amino acid properties were decisive during the finalization of the universal genetic code.

amino acids | genetic code | molecular oxygen | origin of life | redox reactivity

The genetic code embodies the selection and assignment of amino acids (AAs) to be universally employed for protein synthesis. Many aspects of this selection have remained unsettled despite intense intellectual efforts, such that the evolution of the genetic code has occasionally been termed the "universal enigma" of biology (1). Specifically, the choice of the proteinogenic AAs out of a much larger prebiotic and metabolic pool has remained hardly understood, even if a core triad of attributes (size, charge, and hydrophobicity) probably played a major role, especially in the beginning (2). At the same time, theoretical and statistical investigations have clearly suggested that the choice of AAs to be encoded by the standard genetic code was nonrandom and adaptive (2–4).

Attractive candidates to trigger the introduction of new AAs into the genetic code are properties that would enable an improved folding of proteins or higher protein stability (2, 5). Surprisingly, though, investigations on the minimum number of different AA types required to build a foldable polypeptide have concluded that, in general, this number is only about 7–13 (5–9) and may be as low as 3 (10). In various examples of engineered proteins with a reduced AA repertoire, function was maintained (5, 8, 11). Hence, protein folding presently does not provide a ready explanation for the modern expansion of the genetic code beyond the 14-AA stage, and, thus, into the "codon capture" phase of genetic code evolution (12).

Interestingly, the last AAs may have been added substantially later than hitherto assumed. Structural analyses of their aminoacyl-tRNA synthetases indicate that the fixation of tyrosine (Y) and tryptophan (W) in the genetic code only occurred at the time of speciation of the first bacterial lineages (13). Hence, these aminoacyl-tRNA synthetases may not have been present in LUCA, the last universal common ancestor, but rather would have been distributed throughout life at a later time by lateral gene transfer (14). Similarly, the capacity to distinguish the late AA methionine from its genetic code neighbor isoleucine has probably developed only after the radiation of life (15). These observations call for selective factors behind the addition of methionine, Y, and W to the genetic code that have only become relevant to life rather lately.

Quantum chemical properties have not been considered as factors potentially significant for the evolutionary choice of AAs, even though they form the basis of chemical reactivity. For example, out of a list of 566 AA properties assembled in a scholarly database (16), none was referring to characters such as orbital energies or excited states. Hence, we have analyzed the quantum chemistry of the 20 proteinogenic AAs, of numerous abiotic AAs found in meteorites, and of various modern reference molecules. We find that the genetically encoded AAs have become systematically "softer" during early evolution, and we trace this distinctive shift toward higher chemical reactivity of the proteinogenic AAs to the advent of molecular oxygen in the habitat of the early cells.

Results and Discussion

Analyzing the orbital energies of the 20 canonical, proteinogenic AAs and selenocysteine we have found that the energetic distance between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), the HOMO–LUMO gap, exhibited a notable pattern related to the temporal emergence of the AAs in the standard genetic code (Fig. 1): Late

Significance

All life uses the same 20 amino acids, but only 7–13 early amino acids seem to be indispensable to build functional proteins. Thus, what triggered the introduction of the additional amino acids? Employing quantum chemical calculations and biochemical experiments, we find that the additional amino acids have become systematically "softer" (Ralph G. Pearson) over time, more redox-reactive, and more capable of protecting cells from destruction by oxygen free radicals. Hence, it appears that molecular oxygen forced life to incorporate novel amino acids with augmented redox properties into the genetic code. The present study provides a plausible scenario for a more than 80-y-old problem of fundamental biochemistry: Why these 20 amino acids?

Author contributions: M.G. and B.M. conceptualized the project; M.G., P.H., M.I.T., F.-D.I., and B.M. designed research; M.G., P.H., M.I.T., F.-D.I., and B.M. performed research; M.G., P.H., M.I.T., F.-D.I., and B.M. analyzed data; and B.M. wrote the paper.

Conflict of interest statement: Some of the chemical compounds used in this work have been patented by the Max Planck Society, naming author B.M. as one of the inventors (EP 1113795 B1).

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: moosmann@uni-mainz.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1717100115/-/DCSupplemental.

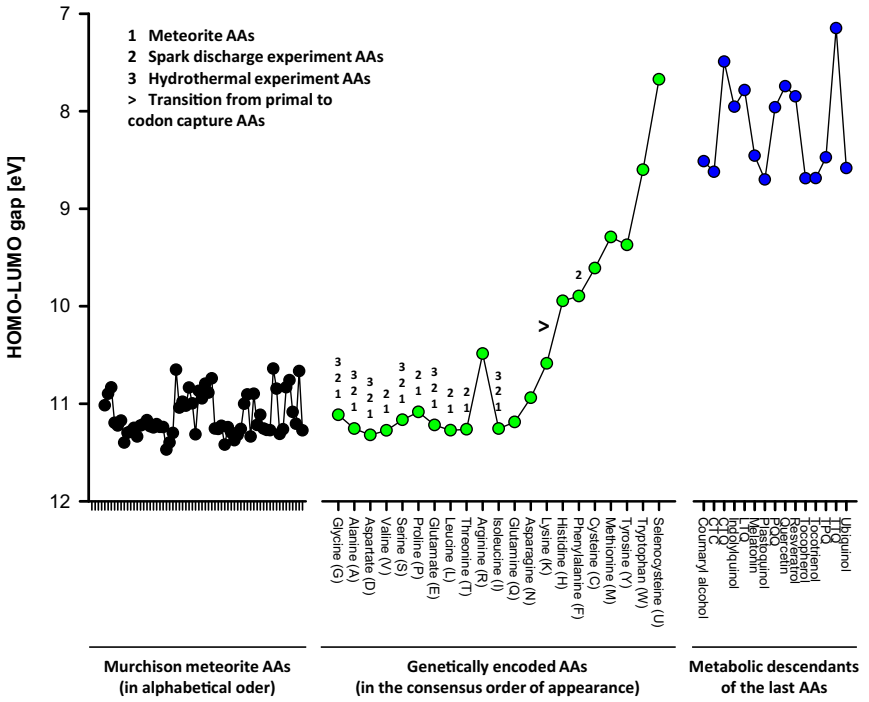

Fig. 1. AA HOMO–LUMO gaps. HOMO–LUMO gaps of 62 Murchison meteorite AAs, 21 genetically encoded AAs, and various metabolic descendants of the shikimate pathway (listed in Table S1) were calculated using semiempirical methods (AM1). The 21 proteinogenic AAs are plotted in the consensus order of their evolutionary appearance according to Trifonov (17), which is in agreement with many single-factor assumptions (e.g., the early introduction of prebiotic AAs, the late introduction of single-codon AAs, or the late introduction of AAs with unusual aminoacyl-tRNA synthetases). Prebiotic AAs identified in cell-free systems (5) and codon capture AAs (12) are indicated. The results of corresponding calculations using a different basis set (PM6) and an ab initio algorithm [Hartree–Fock 6–31+G(d)] are shown in Figs. S1 and S2, respectively. Individual differences notwithstanding, they recapitulate the general principle.

additions to the genetic code (5, 17) have substantially smaller gaps than all early additions, such as the encoded prebiotic AAs found in the Murchison meteorite (18) or recovered from the Urey–Miller experiment (19). Emanating from frontier molecular orbital theory (20), the HOMO–LUMO gap is one of the most widely applicable theoretical predictors of chemical reactivity, particularly reflecting the kinetic stability of a compound toward reactions involving electron transfer or rearrangement (21, 22).

What type of chemical reactivity may explain the evolutionary trend toward smaller HOMO–LUMO gaps? We sought to find reasonable cues in the modern metabolic descendants of the last standard AAs, Y and W, as those descendants might in part represent refinements of the originally selected advantage. Thus, a representative panel of metabolites originating from the shikimate pathway was assembled (Fig. S3). Surprisingly, a rather clear common denominator emerged: The primary function attributable to essentially all examined structures is to catalyze or undergo redox reactions (justified in Table S2). For instance, the role of plastoquinol, ubiquinol, tocopherol, or the quinoid cofactors as committed two-electron and one-electron redox shuttles and redox catalysts is beyond doubt. In terms of their HOMO–LUMO gaps, though, these compounds are quite similar to the last standard AA, W, and the 21st AA, selenocysteine, even if the quinoid cofactors reach absolute orbital energy levels that are unattained by the proteinogenic AAs (Fig. 2). Hence, could the latest AAs indeed be redox cofactors that were integrated into protein structure?

Reviewing the biosynthetic origins of the metabolites of Fig. S3, we noted that all of them involved molecular oxygen (O_2) in at least one step (Table S2). In contrast, all steps of the shikimate pathway, including W and Y biosynthesis, proceed anaerobically. In a framework of Darwinian evolution applied to molecules, W and Y would thus precisely occupy the position of random and initially useless "variants" of the primordial shikimate pathway (toward phenylalanine), whose permanent presence (and introduction into the genetic code) would have only been selected for after the appearance of oxygen in the biosphere.

With this insight, we were able to experimentally test the conclusion emerging from HOMO–LUMO gap analysis that the redox reactivity of the AAs increases with their evolutionary position in the genetic code: We investigated their reactivity toward peroxyl radicals, which are species that must have risen substantially after the cellular emergence of oxygen. The results indicate that, in fact, late AAs are much more reactive toward peroxyl radicals than early AAs (Fig. 3*A*), with the latest additions W and Y being the most effective (Fig. 3*B*). Notably, reactivity under the employed, plausible conditions (aqueous solvent, 1 mM AA, 37 °C) (23) just begins with histidine, the first of the six presumed codon capture AAs: These AAs have probably required a redefinition of already assigned codons (instead of filling up vacancies) (12), particularly calling for selective mechanisms behind their introduction. We have also tested the reactivity of the proteinogenic AAs toward various other radicals. All AAs showing substantial reactivity were newer additions to the genetic code (Fig. S4). However, not all of the late AAs showed reactivity, as a small HOMO–LUMO gap only indicates basic kinetic accessibility of a compound toward electron-rearranging reactants (20); unfavorable transition state energies or other factors may still preclude reactivity, as exemplified in Fig. 3*A* on the inert AA phenylalanine with its high radicalization enthalpy (Table S3).

The biosynthetic selection of a molecule for a specific property does not necessarily imply its introduction into the genetic code. In general, such an introduction indicates the deployment of a molecule to a multitude of macromolecular sites rather than a small number of catalytic centers. In the latter case, the noncovalent binding of a soluble new cofactor might usually be the preferred solution (24). The specific case of selenocysteine demonstrates that an intermediate number of sites (~35 in humans) can suffice to trigger an incomplete integration into the genetic code,

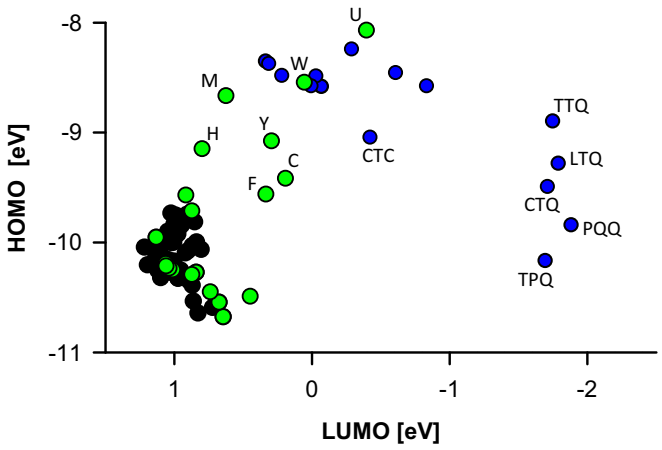

Fig. 2. AA frontier orbital energies. Individual frontier orbital energies are shown of the same compounds as analyzed in Fig. 1. In this representation, small HOMO–LUMO gaps appear in the upper right edge of the graph. All 62 Murchison meteorite AAs and the first 14 genetically encoded AAs cluster in the lower left edge of the graph. Colors and abbreviations are used as in Fig. 1.

whereas a full integration of this AA was not achieved (25). Thus, which topological site of the primordial cell might have demanded the addition of a new, redox-active AA (such as W or Y) as a constituent of proteins in substantial amounts, to fulfill a function that high concentrations of the same AA in soluble form would not have attained? We argue that AA introduction into proteins residing in the lipid bilayer was crucial, for two reasons. First, in many modern cells the lipid membrane is the prime topological site of the cell to be collectively attacked and destroyed by peroxyl radicals in a process called lipid peroxidation, a free radical chain reaction (26, 27). Notably, the toxic effects of high oxygen concentrations are thought to be mediated by such free radical chain reactions (28). However, free AAs are excluded from the lipid bilayer, even those AAs usually classified as lipophilic. Second, many of the assumedly refined metabolic descendants of W and Y that are found ubiquitously primarily differ from their precursors in terms of their sizeable lipid anchors, for example ubiquinol, tocopherol, or plastoquinol (Fig. S3). Their phenolic head groups, however, appear relatively retained, indicating that there was no need for major refinement beyond the original phenolic stage in the three cited examples. Hence, we hypothesized that integration into transmembrane proteins was decisive for the eventual introduction of W and Y into the genetic code.

To explore this idea we have employed differentially lipophilic AA derivatives, as represented by acetylated and dodecanoylated tryptophan ethyl ester shown in Fig. 3*C*. These AA–lipid hybrid molecules were then investigated under "modern" conditions, in classic lipid peroxidation experiments and cellular survival assays under oxidative stress. Exposure of native lipid membranes to Fe^{2+} (10 μM, recycled by 200 μM ascorbate), an abundant component of primordial, low-oxygen, and rising-oxygen habitats (29), led to a pronounced induction of lipid peroxidation, as measured by the formation of malondialdehyde and 8-isoprostane (Fig. 3*D*). Lipophilic derivatives of Y and W (NDo-Y-OEt and NDo-W-OEt) indeed accumulated in membranes (Table S4) and significantly attenuated these radical-mediated reactions, whereas aqueous W and Y derivatives had no effect. To verify this biochemical activity in whole cells, living neurons and fibroblasts were directly exposed to an organic peroxide initiating cellular lipid peroxidation. Fluorescence microscopic images (Fig. 3*E*) and the corresponding quantifications (Fig. 3*F*) demonstrate that supplying cells with lipophilized Y or W prevented peroxide-induced neuronal death, a phenomenon recently discussed in relation to membrane protein oxidation in neurological disease (30). A similarly robust effect was seen in primary fibroblasts (Fig. 3*G*).

We sought to experimentally probe the precision of nature's choice of W as the postulated last redox cofactor to be introduced into the standard genetic code. Hence, we investigated the two most plausible alternatives to W in classical terms (size and shape, charge, and hydrophobicity), which have not been introduced into the code, namely benzofuranyl alanine (BFA) and benzothienyl alanine (BTA). Was the evolutionary choice of W specific for its redox reactivity? As shown in Fig. 4, BFA and BTA were found to be much less reactive toward peroxyl radicals (Fig. 4*B*), and their *N*-dodecanoyl derivatives were essentially unable to prevent lipid peroxidation (Fig. 4*C*) or cell death (Fig. 4*D*), despite the compelling similarity of their side chains to W's side chain in classical terms (Table S5). In what is a rerun of the situation with phenylalanine, BFA's and BTA's relative inertia was unrelated to their HOMO–LUMO gaps but attributable to their very unfavorable radicalization enthalpies (Table S3), which reasserts the role of a small HOMO–LUMO gap as a necessary but insufficient condition for reactivity.

The stepwise increases of oxygen in the biosphere have frequently been cited to represent major, fundamental transitions in the history of life (31–34). However, relatively few examples of specific biochemical changes in response to oxygenation have been presented beyond the expectable consequence of directly O_2-dependent metabolic transformations (32). Here, we present theoretical, metabolic, and experimental evidence that the genetic code itself has been changed in answer to early local oxygenation. Rising oxygen concentrations forced primordial life to deploy increasingly soft molecules [Pearson's absolute chemical hardness η is defined (21) as double of a molecule's HOMO–LUMO gap], to be able to react flexibly to the new challenges imposed by reactive oxygen species that would otherwise disintegrate expensive cellular components designed only for an anoxic habitat (35, 36). With the new AAs, chains of redox reactions starting on protein surfaces could be developed, which would deflect and detoxify radicals in a similar way as present-day organisms (30, 37–39), culminating in the complex one-electron transfer cascades based on W and Y that characterize many modern, oxygen-adapted enzymes (38, 39). For instance, about half of all oxygen-dependent oxidoreductases possess chains of W and Y with three or more members which are stereochemically arranged in such a way as to enable reparative single-electron transfer from the outside (38). Avenues of W and Y seem to play a particularly important role in the control of reactivity of high-potential metal centers in enzymatic oxidation reactions. In case of a prolonged lack of substrate, these activated metal centers are reductively disarmed through chains of Y and W to avoid the looming self-destruction of the enzyme (38).

Intriguingly, a similar strategy of soft AA enrichment was pursued by animal life during the more recent gain in biospheric molecular oxygen about 550 million y ago, at the outset of the Phanerozoic (31, 34). Here, various animal phyla and fungal clades independently changed their mitochondrial genetic code to massively accumulate the soft AA methionine for the purpose of autoprotection (37, 40). Only animal phyla with high metabolic rates and thus high free radical load changed their genetic code assignment (40, 41), and a direct proportionality of aerobic metabolic rate and methionine accumulation (in respiratory chain complex I) has been noted in contemporary animals (41). After all, methionine has long been suspected to represent a genuinely antioxidant AA as it has neither been found to function in enzymatic catalytic cycles nor to be structurally indispensable for its characteristic thioether side chain (37).

Materials and Methods

Reagents. Natural AAs were from Sigma-Aldrich, and *N*-acetylated natural AA derivatives were from Bachem. *N*-dodecanoylated natural AA derivatives were synthesized from free amine precursors as described (42). Enantiopure *L*-(benzofuran-3-yl)alanine (BFA) and *L*-(benzo[*b*]thiophen-3-yl)alanine (BTA) were obtained through chemoenzymatic synthesis starting from benzo[*b*]thiophene and 1-(2-hydroxyphenyl)ethanone as detailed before

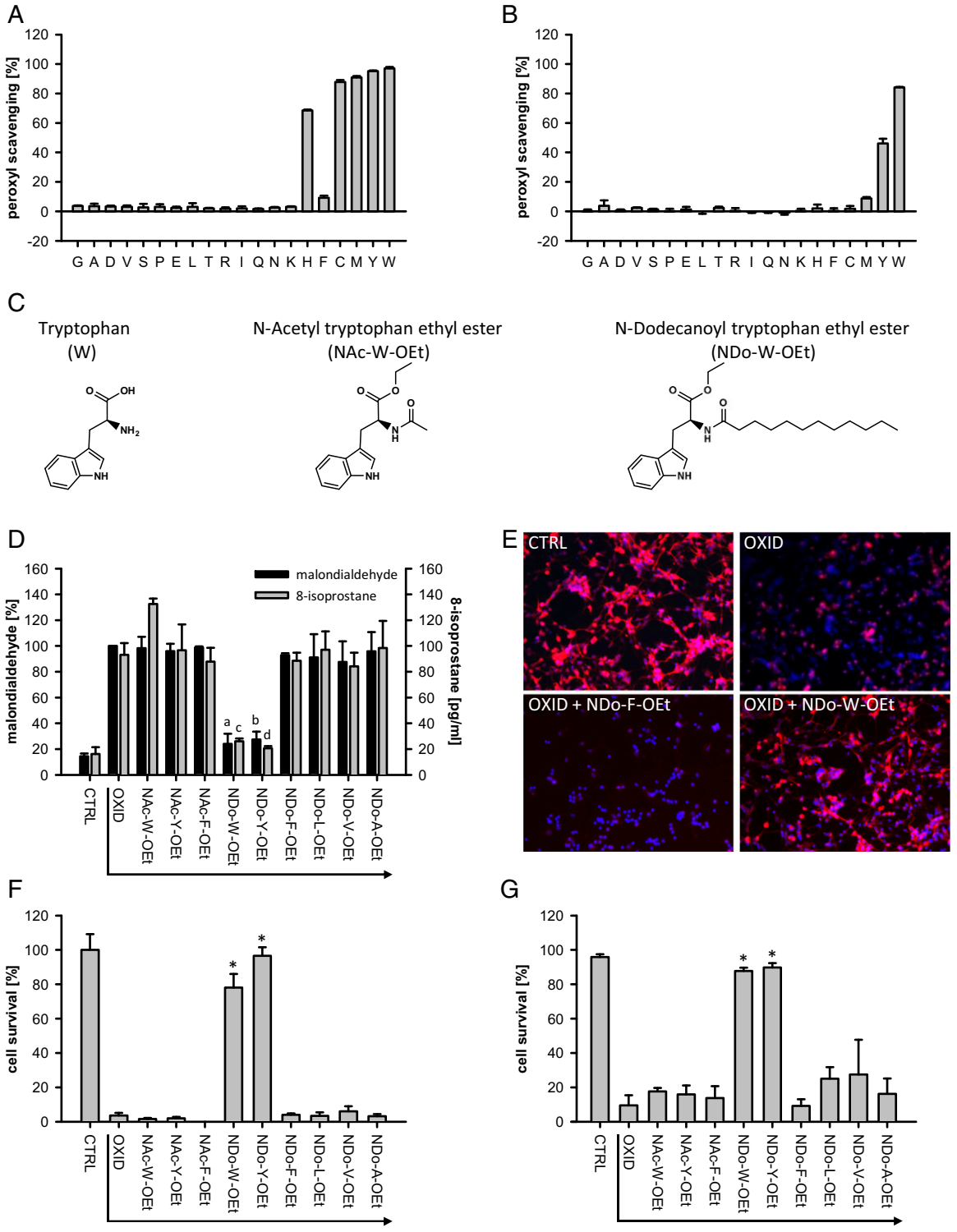

Fig. 3. Redox reactivity of AAs. (*A* and *B*) Effect of the 20 standard AAs in a peroxyl radical scavenging assay. The relative activity of the AAs was tested at ratios 1:3 (*A*) and 1:2,000 (*B*) of AA versus radical initiator ($n = 3$). All AAs below a certain HOMO–LUMO threshold (~10 eV) demonstrated activity under the conditions in *A*, except phenylalanine, whose inertia might be related to its unusually high radicalization enthalpy (Table S3). W and Y were the most effective AAs at high initiator/scavenger ratios (*B*). (*C*) Chemical structures of W, NAc-W-OEt, and NDo-W-OEt. (*D*) Inhibition of lipid peroxidation by acetylated or dodecanoylated AA ethyl esters (100 μM). Ferrous iron-induced lipid peroxidation was monitored by measuring the formation of malondialdehyde ($F_1 = 379$, df = 1; $F_2 = 31$, df = 12; $^{a}P < 0.001$, $^{b}P < 0.001$ versus "OXID" by post hoc test, $n = 3$ for compounds, $n = 9$ for controls) and 8-isoprostane [$F_1 = 35$, df = 1; $F_2 = 11$, df = 12; $^{c}P = 0.002$, $^{d}P = 0.001$ versus "OXID" by post hoc test, $n = 2$ for compounds, $n = 4$ for controls (all from duplicates)]. (*E*) Fluorescence microscopic images of the effect of NDo-W-OEt and NDo-F-OEt (10 μM) on the survival of neurons treated with 100 μM *tert*-butyl hydroperoxide (tBuOOH). Cells were immunostained for microtubule-associated protein 2 (red) indicative of neuronal survival and counterstained with the chromatin marker dye Hoechst 33342 (blue). (*F*) Quantification of cell survival experiments conducted as in *E*. Only lipophilic W and Y derivatives elicited cytoprotective activity against peroxide toxicity ($F_1 = 946$, df = 1; $F_2 = 133$, df = 9; $^{*}P < 0.001$ versus "OXID" by post hoc test, $n = 3$ for compounds, $n = 6$ for controls). (*G*) Survival of fibroblasts treated with 50 μM tBuOOH in the presence of different AA derivatives at 10 μM concentration ($F_1 = 511$, df = 1; $F_2 = 48$, df = 12; $^{*}P < 0.001$ versus "OXID" by post hoc test, $n = 3$ for compounds, $n = 9$ for controls).

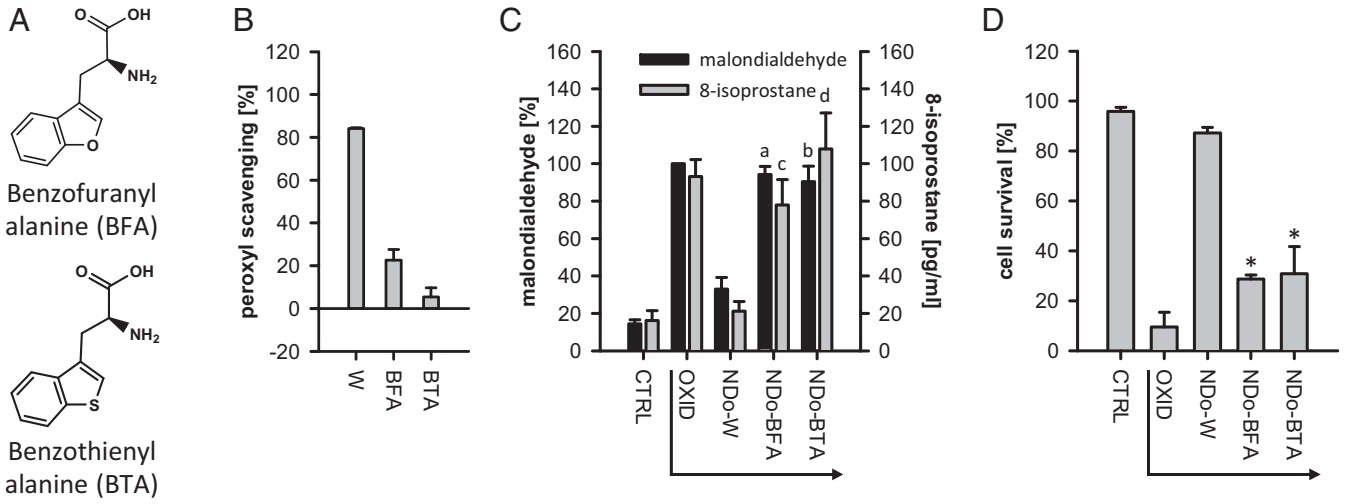

Fig. 4. Specificity of W. (*A*) Chemical structures of BFA and BTA. (*B*) W, BFA, and BTA as scavengers of peroxyl radicals under conditions as in Fig. 3*B* ($n = 3$). (*C*) Inhibition of lipid peroxidation by NDo-W, NDo-BFA, and NDo-BTA under conditions as in Fig. 3*D* [malondialdehyde: $F_1 = 379$, df = 1; $F_2 = 31$, df = 12; $^{a}P < 0.001$, $^{b}P < 0.001$ versus "NDo-W" by post hoc test, $n = 3$ for compounds, $n = 9$ for controls; 8-isoprostane: $F_1 = 35$, df = 1; $F_2 = 11$, df = 12; $^{c}P = 0.014$, $^{d}P = 0.001$ versus "NDo-W" by post hoc test, $n = 2$ for compounds, $n = 4$ for controls (all from duplicates)]. (*D*) Survival of tBuOOH-treated fibroblasts in the presence of NDo-W, NDo-BFA, and NDo-BTA under conditions as in Fig. 3*G* ($F_1 = 511$, df = 1; $F_2 = 48$, df = 12; $^{*}P < 0.001$ versus "NDo-W" by post hoc test, $n = 3$ for compounds, $n = 9$ for controls).

(43). *N*-dodecanoylated BFA and BTA were synthesized as described in *SI Materials and Methods*.

Reagents and solvents were purchased from Sigma-Aldrich if not otherwise stated. In particular, this supplier provided ascorbic acid, 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS), 2,2′-azobis(2-methylpropionamidine) (APPH), *tert*-butyl hydroperoxide, 2,2-diphenyl-1-picrylhydrazyl (DPPH), fluorescein, galvinoxyl, Hoechst 33342, iron(II) sulfate, potassium peroxodisulfate, and 7,7,8,8-tetracyanoquinodimethane (TCNQ).

Compound abbreviations denote the following: NAc-W-OEt is *N*-acetyl tryptophan ethyl ester, NAc-Y-OEt is *N*-acetyl tyrosine ethyl ester, NAc-F-OEt is *N*-acetyl phenylalanine ethyl ester, NDo-W-OEt is *N*-dodecanoyl tryptophan ethyl ester, NDo-Y-OEt is *N*-dodecanoyl tyrosine ethyl ester, NDo-F-OEt is *N*-dodecanoyl phenylalanine ethyl ester, NDo-L-OEt is *N*-dodecanoyl leucine ethyl ester, NDo-V-OEt is *N*-dodecanoyl valine ethyl ester, NDo-A-OEt is *N*-dodecanoyl alanine ethyl ester, NDo-W is *N*-dodecanoyl tryptophan, NDo-BFA is *N*-dodecanoyl benzofuranyl alanine, and NDo-BTA is *N*-dodecanoyl benzothienyl alanine.

Chemical Calculations. Quantum-chemical properties of Murchison meteorite AAs, proteinogenic AAs, modern metabolic descendants of the aromatic AAs, and synthetic tryptophan analogs were calculated using the academic software packages MOPAC (Molecular Orbital Package, releases 2003 and 2009) (44), embedded in the commercial software suites Chem3D 9.0 and ChemBio3D 13.0, respectively, from CambridgeSoft (PerkinElmer), and GAMESS (US) (General Atomic Molecular and Electronic Structure System, release of January 12, 2009) (45), embedded in ChemBio3D 13.0. Chemical structures were assembled, geometry-optimized by energy minimization using a molecular mechanics algorithm (a modified Allinger's MM2 as implemented in Chem3D and ChemBio3D 13.0), and manually inspected for plausibility. Afterward, quantum chemical calculations of electronic energy levels were performed using three different methods.

i) Semiempirical calculations using the MOPAC 2003 AM1 Hamiltonian, via the according MOPAC interface in Chem3D 9.0.
ii) Semiempirical calculations using the MOPAC 2009 PM6 Hamiltonian, via the according MOPAC interface in ChemBio3D 13.0. Compared with AM1, the PM6 basis set provides a parameterization of more chemical elements and a modified treatment of core–core interactions (46).
iii) Ab initio calculations based on the Hartree–Fock method [6-31+G(d) basis set] as implemented in GAMESS (US), via the according GAMESS interface in ChemBio3D 13.0.

Heats of formation and radicalization enthalpies of formation were generally calculated by method *i*. To this end, the standard enthalpy of formation of each structure was calculated before and after the abstraction of a hydrogen radical. The difference of these enthalpies was defined as radicalization enthalpy. If several chemically plausible sites of potential hydrogen radical abstraction were present in a given molecule, all were explored, and the site yielding the lowest radicalization enthalpy was considered further. Dipole moments of the binuclear AA side chains were also calculated using method *i*.

The geometric properties, molecular area, and solvent-excluded volume (in water) as well as molecular hydrophobicity values (logP) were estimated with the ChemPropPro tool set implemented in ChemBio3D 13.0.

UV-visible absorption spectra of W, BFA, and BTA were measured in a spectrophotometer (Beckman Coulter). Compounds were dissolved at 1 mM in PBS, pH 7.4.

Chemical and Biochemical Experiments. Details on the conducted radical assays (peroxyl, TCNQ, galvinoxyl, DPPH, and ABTS radical scavenging) and lipid peroxidation measurements (formation of malondialdehyde and 8-iso prostaglandin $F_{2\alpha}$) can be found in *SI Materials and Methods*.

Cell Culture. Primary cerebellar neurons were prepared from P2 Sprague-Dawley rats and cultivated in Neurobasal medium under a 2% oxygen atmosphere as also present in mammalian brain parenchyma. All other parameters were set as described (30). Analysis of neuronal survival was done by immunocytochemical staining for the neuronal marker protein anti-microtubule-associated protein 2 (MAP-2) following standard protocols (30). Animal procedures were performed in accordance with federal law and the institutional guidelines of the Central Animal Facility of the University of Mainz.

Human lung fibroblasts (type IMR-90; repository number I90) (47) were obtained from the Coriell Institute for Medical Research. The Coriell Institute adheres to Department of Health and Human Services (DHHS) regulations for the protection of human subjects, including informed consent, and distributes samples after submission and review by the Coriell Institutional Review Board of a statement of research intent declaring compliance with DHHS regulations. IMR-90 cells were cultured in DMEM supplemented with 10% heat-inactivated FCS, 1 mM pyruvate, and 1× antibiotic–antimycotic mix (all from Invitrogen). Cells were grown in 100-mm dishes at 37 °C in a humidified atmosphere containing ambient oxygen and 5% CO_2. For the survival experiments, fibroblasts at 80% confluency were seeded into 96-well plates in 0.1 mL medium. After 24 h, the cells were pretreated with the relevant AA derivatives (10 μM) for 5 h, before adding *tert*-butyl hydroperoxide (50 μM). To assess cell survival, the cells were stained with Hoechst 33342 (1 μg/mL) and evaluated by bright-field and fluorescence microscopy.

Statistical Analysis. All data are presented as mean ± SD. Two-way ANOVA was used for all statistical evaluations provided in this work. In general, factor 1 was the presence of a prooxidant or cytotoxin, and factor 2 was the presence of a potential antioxidant or cytoprotector. The indicated levels of significance in the figures (*P* values) result from Holm–Sidak's multiple comparisons procedure executed after the two-way ANOVA under the condition

that overall significance (of $P = 0.001$ or less) was reached in the two-way ANOVA independently for factor 1 and factor 2. Cited n values refer to independent biological or biochemical experiments; technical replicates are referred to as duplicates, triplicates, etc.

ACKNOWLEDGMENTS. We thank C. Behl, U.-G. Maier, and G. Unden for their support. M.I.T. and F.-D.I. thank the Romanian National Authority for Scientific Research, CNCS - UEFISCDI, for financial sponsorship (Project PN-II-ID-PCE-2011-3-0775).

1. Koonin EV, Novozhilov AS (2009) Origin and evolution of the genetic code: The universal enigma. *IUBMB Life* 61:99–111.
2. Lu Y, Freeland SJ (2008) A quantitative investigation of the chemical space surrounding amino acid alphabet formation. *J Theor Biol* 250:349–361.
3. Sengupta S, Higgs PG (2015) Pathways of genetic code evolution in ancient and modern organisms. *J Mol Evol* 80:229–243.
4. Ilardo M, Meringer M, Freeland S, Rasulev B, Cleaves HJ, 2nd (2015) Extraordinarily adaptive properties of the genetically encoded amino acids. *Sci Rep* 5:9414.
5. Longo LM, Blaber M (2012) Protein design at the interface of the pre-biotic and biotic worlds. *Arch Biochem Biophys* 526:16–21.
6. Riddle DS, et al. (1997) Functional rapidly folding proteins from simplified amino acid sequences. *Nat Struct Biol* 4:805–809.
7. Murphy LR, Wallqvist A, Levy RM (2000) Simplified amino acid alphabets for protein fold recognition and implications for folding. *Protein Eng* 13:149–152.
8. Akanuma S, Kigawa T, Yokoyama S (2002) Combinatorial mutagenesis to restrict amino acid usage in an enzyme to a reduced set. *Proc Natl Acad Sci USA* 99: 13549–13553.
9. Fan K, Wang W (2003) What is the minimum number of letters required to fold a protein? *J Mol Biol* 328:921–926.
10. Davidson AR, Lumb KJ, Sauer RT (1995) Cooperatively folded proteins in random sequence libraries. *Nat Struct Biol* 2:856–864.
11. Walter KU, Vamvaca K, Hilvert D (2005) An active enzyme constructed from a 9-amino acid alphabet. *J Biol Chem* 280:37742–37746.
12. Trifonov EN (2004) The triplet code from first principles. *J Biomol Struct Dyn* 22:1–11.
13. Yang XL, et al. (2003) Crystal structures that suggest late development of genetic code components for differentiating aromatic side chains. *Proc Natl Acad Sci USA* 100: 15376–15380.
14. Mukai T, Reynolds NM, Crnković A, Söll D (2017) Bioinformatic analysis reveals archaeal tRNA-Tyr and tRNA-Trp identities in bacteria. *Life (Basel)* 7:8.
15. Jones TE, Ribas de Pouplana L, Alexander RW (2013) Evidence for late resolution of the aux codon box in evolution. *J Biol Chem* 288:19625–19632.
16. Kawashima S, et al. (2008) AAindex: Amino acid index database, progress report 2008. *Nucleic Acids Res* 36:D202–D205.
17. Trifonov EN (2009) The origin of the genetic code and of the earliest oligopeptides. *Res Microbiol* 160:481–486.
18. Cronin JR, Pizzarello S (1983) Amino acids in meteorites. *Adv Space Res* 3:5–18.
19. Miller SL (1953) A production of amino acids under possible primitive earth conditions. *Science* 117:528–529.
20. Fukui K (1982) Role of frontier orbitals in chemical reactions. *Science* 218:747–754.
21. Pearson RG (1986) Absolute electronegativity and hardness correlated with molecular orbital theory. *Proc Natl Acad Sci USA* 83:8440–8441.
22. Aihara J (1999) Reduced HOMO−LUMO gap as an index of kinetic stability for polycyclic aromatic hydrocarbons. *J Phys Chem A* 103:7487–7495.
23. Hren MT, Tice MM, Chamberlain CP (2009) Oxygen and hydrogen isotope evidence for a temperate climate 3.42 billion years ago. *Nature* 462:205–208.
24. Szathmáry E (1993) Coding coenzyme handles: A hypothesis for the origin of the genetic code. *Proc Natl Acad Sci USA* 90:9916–9920.
25. Hatfield DL, Gladyshev VN (2002) How selenium has altered our understanding of the genetic code. *Mol Cell Biol* 22:3565–3576.
26. Gutteridge JM (1995) Lipid peroxidation and antioxidants as biomarkers of tissue damage. *Clin Chem* 41:1819–1828.
27. Yin H, Xu L, Porter NA (2011) Free radical lipid peroxidation: Mechanisms and analysis. *Chem Rev* 111:5944–5972.
28. Fridovich I (1998) Oxygen toxicity: A radical explanation. *J Exp Biol* 201:1203–1209.
29. Ilbert M, Bonnefoy V (2013) Insight into the evolution of the iron oxidation pathways. *Biochim Biophys Acta* 1827:161–175.
30. Hajieva P, Bayatti N, Granold M, Behl C, Moosmann B (2015) Membrane protein oxidation determines neuronal degeneration. *J Neurochem* 133:352–367.
31. Holland HD (2006) The oxygenation of the atmosphere and oceans. *Philos Trans R Soc Lond B Biol Sci* 361:903–915.
32. Raymond J, Segrè D (2006) The effect of oxygen on biochemical networks and the evolution of complex life. *Science* 311:1764–1767.
33. Crowe SA, et al. (2013) Atmospheric oxygenation three billion years ago. *Nature* 501: 535–538.
34. Lyons TW, Reinhard CT, Planavsky NJ (2014) The rise of oxygen in Earth's early ocean and atmosphere. *Nature* 506:307–315.
35. Martin W, Russell MJ (2003) On the origins of cells: A hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos Trans R Soc Lond B Biol Sci* 358:59–83, discussion 83–85.
36. Weiss MC, et al. (2016) The physiology and habitat of the last universal common ancestor. *Nat Microbiol* 1:16116.
37. Levine RL, Mosoni L, Berlett BS, Stadtman ER (1996) Methionine residues as endogenous antioxidants in proteins. *Proc Natl Acad Sci USA* 93:15036–15040.
38. Gray HB, Winkler JR (2015) Hole hopping through tyrosine/tryptophan chains protects proteins from oxidative damage. *Proc Natl Acad Sci USA* 112:10920–10925.
39. Aubert C, Vos MH, Mathis P, Eker AP, Brettel K (2000) Intraprotein radical transfer during photoactivation of DNA photolyase. *Nature* 405:586–590.
40. Bender A, Hajieva P, Moosmann B (2008) Adaptive antioxidant methionine accumulation in respiratory chain complexes explains the use of a deviant genetic code in mitochondria. *Proc Natl Acad Sci USA* 105:16496–16501.
41. Schindeldecker M, Moosmann B (2015) Protein-borne methionine residues as structural antioxidants in mitochondria. *Amino Acids* 47:1421–1432.
42. Fincher TK, Yoo SD, Player MR, Sowell JW, Sr, Michniak BB (1996) In vitro evaluation of a series of N-dodecanoyl-L-amino acid methyl esters as dermal penetration enhancers. *J Pharm Sci* 85:920–923.
43. Podea PV, Tosa MI, Paizs C, Irimie FD (2008) Chemoenzymatic preparation of enantiopure L-benzofuranyl- and L-benzo[b]thiophenyl alanines. *Tetrahedron Asymmetry* 19:500–511.
44. Stewart JJ (2008) MOPAC2009 (Stewart Computational Chemistry, Colorado Springs, CO). Available at OpenMOPAC.net. Accessed May 13, 2016.
45. Schmidt MW, et al. (1993) General atomic and molecular electronic structure system. *J Comput Chem* 14:1347–1363.
46. Stewart JJ (2007) Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J Mol Model* 13:1173–1213.
47. Nichols WW, et al. (1977) Characterization of a new human diploid cell strain, IMR-90. *Science* 196:60–63.
48. Yamamoto Y, Niki E, Kamiya Y (1982) Oxidation of lipids: III. Oxidation of methyl linoleate in solution. *Lipids* 17:870–877.
49. Betigeri S, Thakur A, Raghavan K (2005) Use of 2,2′-azobis(2-amidinopropane) dihydrochloride as a reagent tool for evaluation of oxidative stability of drugs. *Pharm Res* 22:310–317.
50. Cao G, Alessio HM, Cutler RG (1993) Oxygen-radical absorbance capacity assay for antioxidants. *Free Radic Biol Med* 14:303–311.
51. Melby LR, et al. (1962) Substituted quinodimethans. II. Anion-radical derivatives and complexes of 7,7,8,8-tetracyanoquinodimethan. *J Am Chem Soc* 84:3374–3387.
52. Grilj J, Zonca C, Daku LM, Vauthey E (2012) Photophysics of the galvinoxyl free radical revisited. *Phys Chem Chem Phys* 14:6352–6358.
53. Litwinienko G, Ingold KU (2003) Abnormal solvent effects on hydrogen atom abstractions. 1. The reactions of phenols with 2,2-diphenyl-1-picrylhydrazyl (dpph*) in alcohols. *J Org Chem* 68:3433–3438.
54. Re R, et al. (1999) Antioxidant activity applying an improved ABTS radical cation decolorization assay. *Free Radic Biol Med* 26:1231–1237.
55. Moosmann B, Behl C (1999) The antioxidant neuroprotective effects of estrogens and phenolic compounds are independent from their estrogenic properties. *Proc Natl Acad Sci USA* 96:8867–8872.
56. Hajieva P, Mocko JB, Moosmann B, Behl C (2009) Novel imine antioxidants at low nanomolar concentrations protect dopaminergic cells from oxidative neurotoxicity. *J Neurochem* 110:118–132.