Biomedical Physics & Engineering Express

PAPER

Quantum nature of the codon–anticodon interaction

RECEIVED
17 July 2017

REVISED
25 October 2017

ACCEPTED FOR PUBLICATION
9 November 2017

PUBLISHED
8 January 2018

Alexander A Tulub

Argonne National Laboratory, 9700S, Argonne, IL, 60439, United States of America
Centre for Interdisciplinary Computational and Dynamical Analysis, University of Manchester, Oxford Road, Manchester, M13 9PL, Manchester, United Kingdom
Saint-Petersburg State University, University Embankment 7/9, Saint-Petersburg, 199034, Russia

E-mail: atulub@yahoo.co.uk

Keywords: magnesium cofactor, translation, mRNA-tRNA interaction, codon–anticodon communication, spin physics, QM/MM computations

Abstract

Despite impressive progress in molecular biology and genetics over the last decades, the key question is still unanswered—why the genetic code is triplet and what information it keeps and transfers over translation. Unexpectedly, the answer comes not from biology but from spin physics and computational quantum chemistry. The study aims to shed light on how mRNA sees the right tRNA and what sort of communication ensures the codon–anticodon recognition. The data rests on the quantum mechanical/molecular mechanical quantum chemistry computations of the mRNA-tRNA fragments differing in 61 codon–anticodon nucleotide triplets coding 20 canonical amino acids.

1. Introduction

In molecular biology and genetics, *translation* is the second major step in gene expression [1]. Translation proceeds on ribosomes where the genetic code information, initially stored in DNA, transfers from mRNA to a loaded tRNA (tRNA with the attached amino acid), figure 1. The deciphering of the genetic code occurs in the codon (mRNA)-anticodon(tRNA) region [2]. The hydrogen bonding in this region obeys the Watson–Crick pairing rule [3]. Each codon and anticodon consists of three nucleotides. tRNA serves as a bridge between the codon and the right amino acid. What is outlined is a proved fact confirmed by numerous experiments [4]. What is still unanswered is summarized as follows:

(i) Communication between the codon and anticodon does not proceed without Mg cations (cofactors) bound respectively to the codon and anticodon trinucleotides, figure 1.

(ii) There is no understanding why a sequence of three nucleotides (not two, four, five etc) is required for each codon and anticodon.

(iii) The 'wobble' effect (the third nucleotide in the codon–anticodon interaction is less discriminatory for determining the right amino acid [3]) is still unclear.

(iv) What carrier ensures the genetic code operation?

(v) How this carrier is coupled with the codon nucleotide trinity?

The paper aims to elucidate the outlined issues. Conceptually, the research stems from the ideas of spin physics: electron spin flip [5], triplet(T)/singlet(S) potential energy surfaces (PESs) and their intersections [6–8], and Berry's spin phase flip [9], section 3. The named effects appear in DNA molecule upon the action of Mg cations [10–12]. This is seemingly valid for mRNA-tRNA interactions [10, 11]. The results rest on the quantum mechanical (QM)/molecular mechanical (MM) computations for 61 codon–anticodon possible triplets coding 20 canonical amino acids.

2. Modeling and computations

Translation is a highly complicated process to be studied quantum chemically. So, drastic simplifications are required. Meanwhile, the typical features of the process should be preserved. With this in mind, we reduce the mRNA to three parts—the codon (its composition spans over 61 triplets) in the center and twenty identical, invariable by their composition, triplets (AAA) on its both sides. The tRNA is considered without cuts, including the attached amino acid

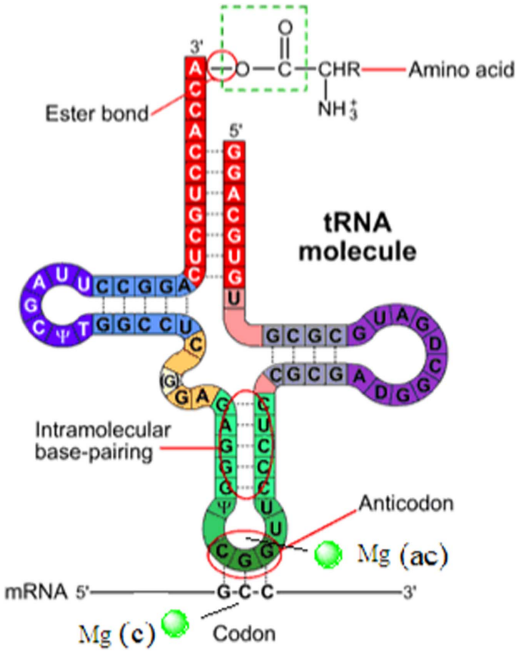


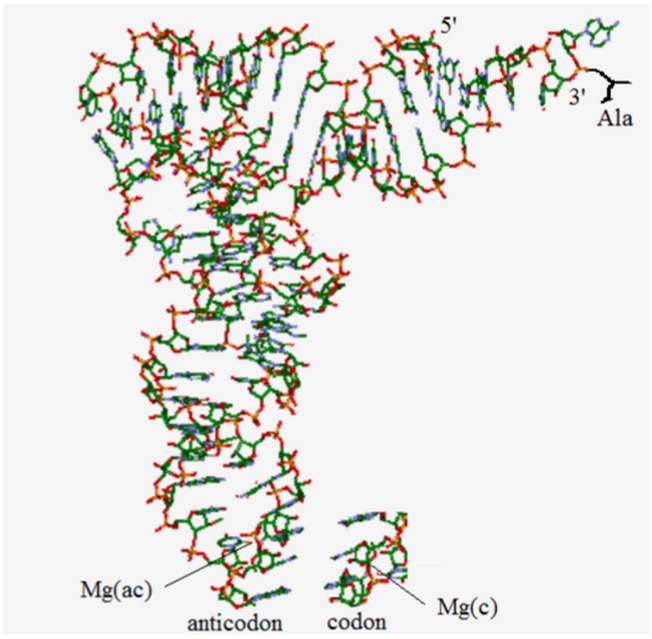
Figure 1. 2D structure of the tRNA-mRNA complex with two Mg cofactors (water molecules not shown) bound to the codon (mRNA), Mg(c), and anticodon (tRNA), Mg(ac).

(in some cases, its nature was intentionally varied, see below). The ribosome is excluded from consideration. Actually, the system under study appears as that in figure 1.

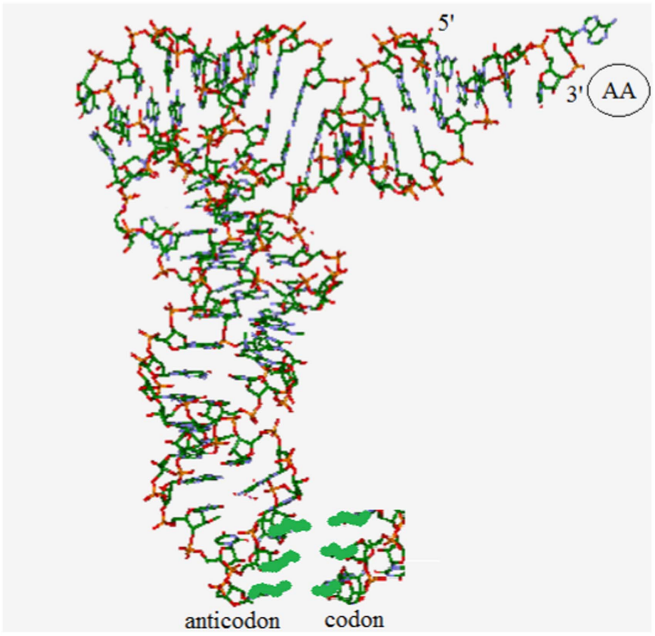
The QM/MM is rather the only one applicable method for treating huge biological molecules. Any QM/MM method consists of two parts—the quantum mechanical (QM) part and the MM part. The QM part is computationally demanding and time consuming. So, it covers the most significant parts (sufficiently small) of the molecule. In our case, the QM part includes the codon and anticodon with two magnesium cofactors and the attached amino acid (AA). The rest part of the system (the twenty nucleotide triplets, AAA, on both sides of the codon and the huge part of tRNA between the anticodon and the attached amino acid, figure 1) is considered MM (this excludes any ambiguities with the terminal triplets (e.g., addition of His) and non-compensated charges over all the triplets from diester bonds because the QM part does not include these specifics into consideration; on the contrary, the addition of counter-ions into the MM part leads to unexpected distortions of the lengthy AAA chains, which are still unbound to any template in our model). The QM part is treated with the Gromacs DFT:B3LYP (6-311G** basis set) code, version 6.52, and the MM part with the AMBER-8.5 code included into the same Gromacs [10, 11, 13]. The computations rest on the automated fragmentation approach (AF-QM/MM) [13], which ensures the QM/MM boundaries. The use of B3LYP with polarization functions (specifically, they are necessary to treat long-range interactions like hydrogen bonding) is justified because of its numerous tests on biological molecules of different length, composition and structure. The initial geometry of the loaded tRNA (61 fragments) and the mRNA fragment (the codon (61 units) and the twenty AAA triplets, see above) are from the Protein Data Bank [14]. The computations assume the geometry optimization procedure both for the QM part, including the cofactors, and the MM part. The approach between the mRNA and the loaded tRNA fragments (the codon–anticodon region) proceeds in a step-by-step manner (each step is 0.05 Å; the initial separation distance between the codon and anticodon is 4 Å) to locate the minimum value of the system (mRNA fragment + loaded tRNA fragments) total energy, E. In the case of Ala, Thr, Tyr and Pro loaded tRNAs, we made round artificial replacements (Pro → Thr, Tyr → Ala etc) leaving the codon–anticodon complementary triplets invariable. The purpose of these replacements is clarified in the next section. The $Mg^{2+}(H_2O)_2$ acts as a magnesium cofactor. Its choice is due to the fact that the magnesium cofactor in its active state forms two bonds with water molecules, leaving the two (or one) unused valences to form bonds/bond with two or one oxygen of the nucleotide diester bond (the inert form assumes six bonds with water molecules, which leaves no chance of making bonds with the diester) [10]. The two cofactors (they are right two because the addition of other magnesium cofactors to other possible diesters in the QM part does not change the picture, see [7, 11, 12]; in addition, the tRNA-mRNA structure shows two magnesiums built into the codon–anticodon region: one magnesium is bound to the anticodon, the other is bound to the codon [14]) are positioned, respectively, on the codon (Mg(c)) and anticodon (Mg(ac)), figure 1 (water molecules are not shown), and are allowed to find the 'best' nucleotide to bind over the mRNA-tRNA optimization ensuring the minimum of E^{tot}. Figuratively, each cofactor is allowed to drift over the QM part looking for energetically more favorable position to occupy (within a set of available oxygens). In practice, this is achieved through step-by-step displacements (0.05 Å) of the magnesium cofactor over a set ($m = 1, 2, 3$) of manifolds (each manifold is equal to a single nucleotide) covering the QM part.

3. Results and discussion

Upon optimization, the Mg(c) and Mg(ac) make strong bonds with negatively charged oxygens (the Mg–O distance is 1.89 Å (codon) and 1.91 Å (anticodon)) and weak bonds with the other oxygens (the Mg–O distance is 2.28 Å (codon) and 2.29 Å (anticodon)) of the phosphodiester fragment. This fragment links the first and the second nucleotides (codon), and the third and the second nucleotides (anticodon), figure 2(a). What we have is identical to what we have seen in the Mg-bound trinucleotides of DNA [10, 11].



(a)



(b)

Figure 2. (a) The optimized structure of the Ala loaded tRNA with the complementary bound GCC codon (the AAA triplets on both sides of the codon not shown). (b) The optimized structure of tRNA loaded with the wrong amino acid (AA) = Tyr. As before (figure 2), the codon is GCC and the anticodon is CGG. The codon and anticodon triplets reveal the T_+ symmetries with the uncompensated spin density (shown in green) spreading over the triplets.

The outlined configuration makes Mg cofactors singly charged ($q = +1$) near the crossing/intersection point (see below) with an unpaired spin each—'up' (↑) or 'down' (↓). As soon as the Mg(c) and Mg(ac) spins produced, their counterpart spins arise sequentially—this time on the codon and anticodon. The computations show that upon approaching the mRNA fragment to the tRNA (the contact region is the codon–anticodon) the Mg(c)-codon fragment reveals the (↑↑) configuration (T_+), while the Mg(as)-anticodon fragment reveals either T_- configuration (↓↓) or T_+ configuration (T_+, T_- notations for triplet states with spins up and spins down were introduced into spin chemistry by Turro [15]). *The choice between the T_+ and T_- anticodon strongly depends on what type of loaded RNA interacts with the codon. If that is the right loaded tRNA* (*the amino acid corresponds to the codon,* figures 1, 2(a)), *then we have the T_+ codon and the T_-*

anticodon; if not (*the amino acid does not correspond to the codon*), *we have the* T_+ *codon and the* T_+ *anticodon* (figure 2(b)). The statement finds its confirmation. When we begin to approach the mRNA fragment and the loaded tRNA fragment, their PESs reveal a crossing point at the distance of 3.4 Å between the codon and anticodon, figures 3(a), (b) (2D PESs cuts). Right at this point we observe the (T_+)–(T_+) or the (T_+)–(T_-) spin configuration. The (T_+)–(T_+) configuration corresponds to the wrongly attached amino acid (the attachment is made artificially before starting the optimization procedure). Figure 2(b) shows the spin density distribution (the (T_+)–(T_+) configuration, AA = Tyr) over the codon and anticodon. The (T_+)–(T_-) configuration corresponds to the amino acid attached in the right way, AA = Ala, figure 2(a) (spins are totally compensated, and no spin density is revealed). The (T_+)–(T_+) configuration arises each time when we replace the right amino acid by the wrong one (Ala → Tyr, Thr, Ser, Val; Pro → Thr, Ser, Arg, Val etc) leaving the codon–anticodon paired right (GCC-CGG; CCC-GGG etc). Earlier it was proved that the nucleotide triplets with the identical spin orientation experience the repulsion [12]. The same we observe in our case with the two PESs (the crossing point) of identical spin configuration, figure 2(b) (AA = Tyr),—the value of E at a 4.0 Å distance is lower than that at 3.4 Å by 2.4 kcal mol^{-1}. (Our artificial approach between the tRNA and mRNA is repulsive!) The result is of high importance. It indicates that *the wrongly loaded tRNA never binds to the mRNA, even if the codon and anticodon obey the Watson–Crick pairing rule. The decision of making or not making the pairing is made at a distance* (3.4 Å) *exceeding the distance of hydrogen bonded nucleotides by* 0.5 Å [1]. On the contrary, when the amino acid is right, the repulsion at 3.4 Å does not occur. Moreover, the two PESs reveal the attraction which leads to "plunging" the upper PES (curve 2) into the bottom PES (curve 1) with the appearance of the local energy minimum (E_{min}), the value of Δ, figures 3(c), (d). This minimum is a result of spin pairing which arises as a response to consecutive hydrogen bonding between the complementary nucleotides on the codon and anticodon (the other two spins form the T_- configuration outside the crossing region, figures 3(c), (d)). This is clearly seen in the QM/MM experiment when the QM part initially covers the first complementary nucleotides on the codon and anticodon (the other two are treated MM), then the first and second complementary nucleotides (the rest one is treated MM) and finally all the three complementary nucleotides. It is important to stress that *the spin pairing, corresponding to the* E_{min} (figures 3(c), (d)), *occurs only upon reaching hydrogen bond pairing between all the three nucleotides on the codon and anticodon*. The result is identical to that we observed in DNA paired trinucleotide loops [11]. Its explanation rests on the Berry spin flip [9]. The flip is the manifestation of the

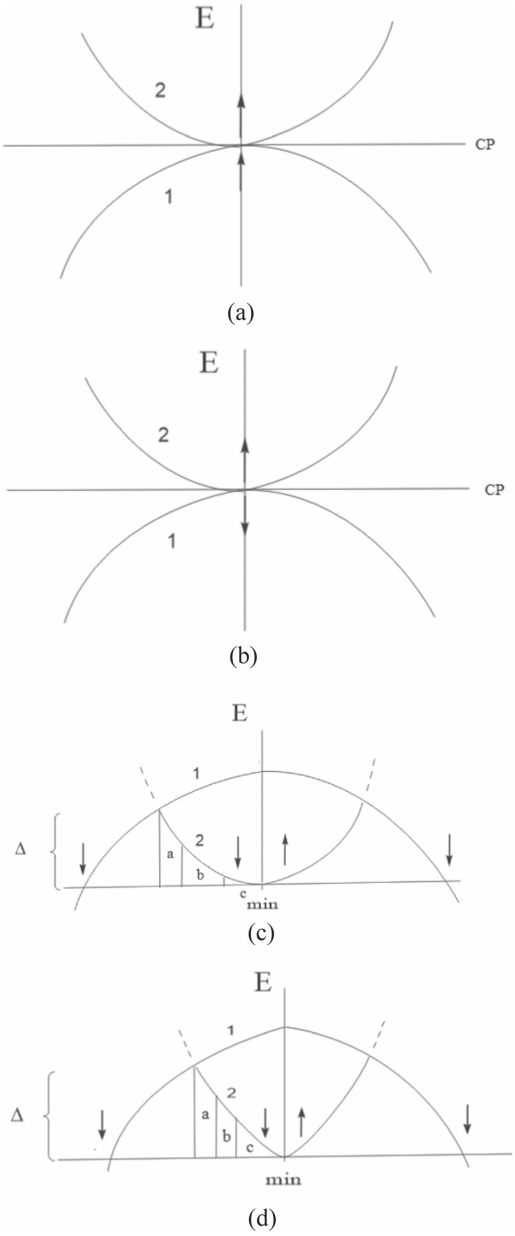


Figure 3. (a) The two PESs (2 corresponds to the mRNA fragment; 1 corresponds to the loaded tRNA fragment) revealing the crossing point. The 'up' arrows indicate the T_+ and T_+ spin configurations belonging to 1 (GCC codon) and 2 (CGG anticodon, AA = Tyr (the wrong amino acid)) PESs. CP is the crossing point. (b) The two PESs (2 corresponds to the mRNA fragment); 1 corresponds to the loaded tRNA fragment revealing the crossing point. The 'up' and 'down' arrows indicate respectively the T_+ (1 PES) and T_- (2 PES) spin configurations belonging to 1 (GCC codon) and 2 (CGG anticodon, AA = Ala (the right amino acid)) PESs. CP is the crossing point. (c) The crossing between the mRNA (curve 2) and the Ala loaded tRNA (curve 1) PESs. The codon and anticodon are hydrogen bonded at the distance of 2.9 Å. a, b, c are the sectors corresponding to the contribution of the first, second and the third hydrogen paired nucleotide (codon (GCC)-anticodon(CGG)) into the energy E. Δ corresponds the energy (12.4 kcal mol^{-1}) between the min and the 2–1 crossing. (d) The crossing between the mRNA (curve 2) and the Trp loaded tRNA (curve 1) PESs. The codon and anticodon are hydrogen bonded at the distance of 2.8 Å. a, b, c are the sectors corresponding to the contribution of the first (U), second (G) and the third (G) hydrogen paired nucleotide (codon(UGG)-anticodon(ACC)) into the energy E. Δ corresponds the energy (12.6 kcal mol^{-1}) between the min and the 2–1 crossing.

non-vanishing geometric phase (Berry's phase) [9], arising in the exponential term of the wave function representing the behavior of two coupled electrons (our case; in general, the number of electrons could be any). When the coupled electrons move around each other (topologically, this is equivalent to particle interchanging) over the trinucleotide area (the QM part), their statistics (phase), $\Delta\theta$, accumulates. The accumulation proceeds in steps ($\Delta\theta_i = \pi/3$, $i = 1, 2, 3$), each corresponding to the statistics 'jump'. The statistics jumps are no more than a response to the vector field (A_i, $i = 1, 2, 3$) change upon the electron 'traveling' over the trinucleotide loop (first nucleotide → second nucleotide → third nucleotide). Upon reaching the value $\Delta\theta = \pi$, the initial direction of spin changes to its opposite one (spin flip!) [11]. *When this value is reached* ($\Delta\theta = \sum_i \Delta\theta_i = \pi$, $i = 1, 2, 3$—the number of paired nucleotides; each $\Delta\theta_i = \pi/3$; SU(2) rotation space), *the electrons, which are spread over the codon and anticodon, obtain the oppositely directed spins, and the pairing occurs; if not, the pairing does not occur because the Berry flip is not reached* [10, 11].

The Ala amino acid is coded by four codons—GCU, GCC, GCA and GCG differing in the nature of the third nucleotide. Figure 3(c) shows that the contribution of these three nucleotides into the value of $\Delta = 12.4$ kcal mol^{-1} is uneven: 4.7, 5.9 and 1.8 kcal mol^{-1}, respectively (a, b, c sectors). These values is a result of sequential enlarging of the QM part (codon + anticodon) from one nucleotide to three (see above). The minor contribution is associated with the third nucleotide—1.8 kcal mol^{-1}. *This is right the wobble effect. Its explanation, from the computational point of view, is in a relatively small contribution of the third nucleotide of codon into the system stabilization*. When the amino acid is coded by a single nucleotide triplet, like with the Trp (UGG codon), figure 3(d), the contributions from each nucleotide are even: 4.2, 4.3, 4.1 kcal mol^{-1} ($\Delta = 12.6$ kcal mol^{-1}). As a result, *there is no wobble effect when each nucleotide of the codon contributes evenly into the system stabilization*. The PESs corresponding to the amino acids coded by two or three codons resembles those shown in figures 3(c), (d). The third nucleotide, meanwhile, is still of lesser importance than the first and second ones. The results are valid for the set of 61 codons (61 mRNAs) and 61 anticodons (61 loaded tRNAs).

QM states that spin is observable just along one axis in the Cartesian frame (z, x, y). If we fix such a reference frame to an arbitrary mRNA-tRNA, the other frames fixed to other mRNA-tRNAs experience a rotation relative to the reference one. This occurs because of changing the direction of spin polarization in the Pauli spin space [5] (the correspondence between the amino acid nature and the spin polarization angle is given in [11], figure 6). The full spin turn in SU(2) corresponds to the value of 4π with each π producing a spin flip. Each flip corresponds to the sum of three $\pi/3$ phases associated with each nucleotide in the codon/anticodon, see above. *Four flips and three phases per a flip generate the famous number of* $64 = 4^3$ *codons coding twenty canonical amino acids plus three stop-codons. Back then, this number appeared combinatorially—four possible nucleotides with three possible* permutations [16, 17]. *Now this number obtains its physical meaning—it is associated with four spin flips and three phases per a flip*.

4. Concluding remarks

The results can be summarized as follows.

(i) Mg cofactors are integral parts of the codon–anticodon communication over translation. Their role is in the production of unpaired electrons in the codon–anticodon region. This is a specific feature of Mg differing it from other doubly-charged analogous cations like Ca and Zn [12].

(ii) The wobble effect is thought as a minor energy contribution into the stabilization energy from the third nucleotide of the codon. Depending on the form of the PES, figures 3(c), (d), we observe the degree of degeneracy: from four to zero.

(iii) Spin appears as a carrier of information stored in the codon. The codon spin consistency with the anticodon spin, (T_+)–(T_-) spin states, ensures the right recognition. The inconsistency, (T_+)–(T_+) spin states, stops the codon–anticodon pairing. Spin states 'feel' each other at distances exceeding those of hydrogen bonding.

(iv) The codon–anticodon recognition requires that both nucleotide sequences come in trinity. The trinity ensures that spin projections from the codon and anticodon on the chosen axis come up in consistency. With no trinity, the electron phase lies between π and zero (in our case this is the value of $\pi/3$ for each nucleotide pair). These phases correspond to 'anyons', particles which are not bosons or fermions. Anyons are widely discussed in the theory of superconductivity and in the theory of braids [18]. Following the concept of anyons, *each nucleotide pair in each nucleotide trinity might be thought as a single anyon*.

(v) The combinatorially treated value of 64 codons coding twenty canonical amino acids plus three stop-codons finds its physical explanation. The value is the manifestation of SU(2) symmetry ensuring the recognition process between the codon and anticodon.

Acknowledgments

The study is sponsored by the NASA Ecology grant NASA-05789-2015-20-A.

ORCID iDs

Alexander A Tulub https://orcid.org/0000-0003-2042-4532

References

[1] Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K and Walter P 2016 *Molecular Biology of the Cell* 6th edn (New York: Garland Science)
[2] Schleif R 1993 *Genetics and Molecular Biology* (Baltimore-London: John Hopkins University Press)
[3] Witts R N, Hopson E C, Koballa D E, van Boening T A, Hopkins N H, Patterson E V and Nagan M C 2013 *J. Phys. Chem.* B **117** 7489–97
[4] Maniatis T 2012 *Nat. Med.* **18** 1499–502
[5] Su D R 1992 *Chin. J. Phys.* **30** 357–87
[6] Atchity G J, Xantheas S S and Ruedenberg K 1991 *J. Chem. Phys.* **95** 1862
[7] Tulub A A 2006 *Phys. Chem. Chem. Phys.* **8** 5368–9
[8] Tulub A A 2007 *Chem. Phys. Lett.* **436** 258–65
[9] Berry M V 1984 *Proc. R. Soc.* A **392** 45–57
[10] Tulub A A and Stefanov V E 2017 *J. Theor. Biol.* **416** 144–8
[11] Tulub A A 2016 *RSC Adv.* **6** 81666
[12] Tulub A A 2016 *Arch. Biochem. Biophys.* **607** 44–6
[13] He X, Zhu T, Wang X, Liu J and Zhang J Z 2014 *Acc. Chem. Res.* **47** 2748–57
[14] Berman H M, Westbrook J, Feng Z, Gilliland G, Bhat T N, Weissig H, Shindyalov I N and Bourne P E 2000 *Nucleic Acids Res.* **28** 235–42
[15] Turro N J, Ramamurthy V and Scaiano J C 2010 *Modern Molecular Photochemistry of Organic Molecules* (Mill Valley, CA: University Science Books)
[16] Taylor F J and Coates D 1989 *BioSystems* **22** 177–87
[17] Koonin E V and Novozhilov A S 2009 *IUBMB Life* **61** 99–111
[18] Balachandran A P, Ercolessi E, Morandi G and Sriva A M 2016 *Hubbard Model and Anyon Supercoductivity* (Singapore-Hong-Kong: World Scientific) Sec. Ed.