



UNIwersYTET
WARMIŃSKO-MAZURSKI
W OLSZTYNIE

Maciej Pyrka

**Właściwości fizyko-chemiczne wybranych 8-azapuryn oraz
specyficzna rybozylacja 2,6 – diamino-azapuryny przez
fosforylaze nukleozydów purynowych**

Physico-chemical properties of selected 8-azapurines
and specific ribosylation of 2,6 - diamino-
azapurine by purine nucleoside phosphorylase

Praca doktorska wykonana w Katedrze Fizyki i Biofizyki na Wydziale Nauki o Żywności
Promotor: dr hab. Maciej Maciejczyk, prof. UWM

Łódź, 2020

Podziękowania

Pragnę serdecznie podziękować **Panu Profesorowi Maciejowi Maciejczykowi** za opiekę naukową, poświęcony czas, przekazaną wiedzę, możliwość czerpania wzorców w pracy naukowej oraz wyrozumiałość.

Serdecznie dziękuję **Panu Profesorowi Zbigniewowi Wieczorkowi**, kierownikowi Katedry Fizyki i Biofizyki, za wsparcie umożliwiające rozwój naukowy.

**Badania prowadzone w ramach niniejszej pracy doktorskiej
były finansowane z następujących źródeł:**

- Grantu wewnętrznego Wydziału Nauki o Żywności Uniwersytetu Warmińsko-Mazurskiego dla młodych naukowców oraz uczestników studiów doktoranckich w latach 2014-2016
- Środków statutowych Katedry Fizyki i Biofizyki Uniwersytetu Warmińsko-Mazurskiego
- Grantu obliczeniowego WCSS (Wrocławskie Centrum Sieciowo-Superkomputerowe)

Spis treści

I.	Spis publikacji wchodzących w skład rozprawy doktorskiej	5
II.	Omówienie celu naukowego i uzyskanych wyników	6
III.	Streszczenie w języku polskim	20
IV.	Streszczenie w języku angielskim	22
V.	Dorobek naukowy	24
VI.	Oświadczenia współautorów prac	27
VII.	Kopie publikacji składających się na rozprawę doktorską	33

I. Spis publikacji wchodzących w skład rozprawy doktorskiej

Maciej Pyrka, Maciej Maciejczyk*; Theoretical study of tautomeric equilibria of 2,6-diamino-8-azapurine and 8-aza-iso-Guanine; *Chemical Physics Letters*, 2015, **627**, 30-35.

Praca oryginalna

Punkty MNiSW: **25** (2015), **70** (2019)

Impact Factor: **1,897**

Maciej Pyrka, Maciej Maciejczyk*; Theoretical investigations of tautomeric equilibrium of 9-methyl-8-aza-iso-Guanine and its electrostatic properties; *Computational and Theoretical Chemistry*, 2016, **1091**, 1-7.

Praca oryginalna

Punkty MNiSW: **20** (2016), **40** (2019)

Impact Factor: **1,403** (2016),

Maciej Pyrka, Maciej Maciejczyk*; Why purine nucleoside phosphorylase ribosylates 2,6-diamino-8-azapurine in non-canonical positions? A molecular modeling study; *Journal of Chemical Information and Modeling*, 2020, **1091**, 1-7.

Praca oryginalna

Punkty MNiSW: **100** (2019)

Impact Factor: **3,966** (2018)

(*) autor korespondencyjny

Sumaryczny IF prac wchodzących w skład rozprawy doktorskiej: 7,266 (210 pkt MNiSW)

II. Omówienie celu naukowego oraz uzyskanych wyników

Wprowadzenie

Równowaga tautomeryczna zasad nukleinowych jest jednym z kluczowych czynników odpowiadających za poprawne rozpoznawanie odpowiadających sobie zasad w kwasach nukleinowych [1], a jej zaburzenie ma duży wpływ na biologiczne procesy, ponieważ może prowadzić do mutacji genetycznych [2]. Proces tautomerizacji puryn oraz zasad nukleinowych został szeroko zbadany zarówno przy użyciu metod teoretycznych jak i eksperymentalnych [3-11]. Rozwój alternatywnego kodu genetycznego (expanded genetic code) zrodził naturalne pytanie o równowagę tautomeryczną jego nowych form – izo-guaniny oraz izo-cytozyny [12]. Problem ten był szczególnie interesujący, ponieważ dane eksperymentalne ujawniły możliwość mutacji izo-cytosyna->tymina [13]. Zsyntetyzowana pochodna guaniny (Gua) – aza-guanina (azaGua) – oraz pochodna izo-guaniny (izoGua) – 8-aza-izo-guanina (z8izoGua) – zostały rozpoznane jako obiecujące sondy fluorescencyjne [14,15], które mogą znaleźć zastosowanie w badaniach ważnych biologicznych procesów. Zamiana węgla C8 na azot nie wpływa bezpośrednio na interfejs Watsona-Cricka i Hoogsteena, chociaż może prowadzić do zmiany równowagi tautomerycznej nowych związków w porównaniu do naturalnych puryn, w tym również do Gua oraz IzoGua.

Oprócz konwersji C8->N8, warto również zbadać ewentualny wpływ metylacji na równowagę tautomeryczną pochodnych puryn. Metylowana izoGua została zbadana przy użyciu metod chemii kwantowej [16], jednak nie pod kątem wpływu metylacji w pozycji 9 na równowagę tautomeryczną. W ramach tych badań grupa metylowa została umieszczona w pozycji nr 9 izoGua w celu imitowania jej obecności w łańcuchu polinukleotydowym. Metylacja umożliwia także wskazanie lokalizacji protonu na pierścieniu triazolowym (w niektórych przypadkach także pierścienia pirymidynowego) azapuryn za pomocą wyznaczonych eksperymentalnie widm absorpcji, co zastosowano min. dla azaGua oraz z8izoGua [15]. Co istotne, metylacja tych związków nie doprowadziła do utraty właściwości fluorescencyjnych.

Bardzo interesującym białkiem, które jako substraty wykorzystuje naturalne puryny oraz ich analogi, jest fosforylaza nukleozydów purynowych (PNP). PNP jest enzymem, który katalizuje odwracalny proces konwersji (rybozylacja i fosforoliza) między zasadami nukleinowymi (purynami) i ich nukleozydami. PNP odgrywa ważną rolę w metabolizmie nukleotydów, ponieważ uczestniczy w zapasowym szlaku metabolicznym syntezy nukleotydów, który wykorzystuje zasady nukleinowe i nukleozydy dostępne w komórce. Jest to alternatywna droga do bardziej powszechnego, ale energetycznie droższego procesu syntezy *de novo*

[17]. Takie właściwości biochemiczne umożliwiają wykorzystanie PNP w procesach farmakologicznych, medycznych i praktycznych. Jedną z negatywnych konsekwencji działania enzymu jest fosforoliza leków nukleozydowych, którą można złagodzić poprzez zastosowanie odpowiednich inhibitorów [18]. Z drugiej strony fosforylaza mogą przyczyniać się do aktywacji proleków, które mogą być nukleozydami lub zasadami nukleinowymi [19]. Niektóre z enzymów PNP można stosować w terapii genowej raka, w której cytotoksyczny kwas nukleinowy uwalnia się w wyniku fosforolizy toksycznych nukleozydów [20]. Wykazano, że niedobór lub brak aktywności PNP prowadzi do dysfunkcji komórek T i powoduje zmniejszoną odporność komórek. PNP może być również stosowane jako lek immunosupresyjny, zapobiegający odrzuceniu przeszczepu, lek na raka powodującego nadprodukcję komórek T i lek na choroby autoimmunologiczne, takie jak dna moczanowa, reumatoidalne zapalenie stawów, łuszczyca, stwardnienie rozsiane [21-23]. Z drugiej strony aktywność enzymatyczna PNP może być wykorzystywana do syntezy specyficznych nukleozydów, co stanowi alternatywną metodę dla syntezy chemicznej [24-27].

Dzięki temu, że 8-azapuryny wykazują mierzalną emisję fluorescencji, to mogą znaleźć zastosowanie w badaniu kinetyki procesu wiązania do PNP [15]. Fluorescencja 8-azapuryn w tym azaGua zależy także od pH i dlatego może być stosowana do badania stanów jonizacji zasad nukleinowych w ustrukturyzowanym RNA [28]. Ostatnie badania eksperymentalne na 8-azapurynach pokazują, że PNP może katalizować rybozylację w różnych pozycjach pierścienia triazolowego, niekoniecznie w pozycji kanonicznej nr 9 [29]. Przyczyna tej różnorodności może wynikać z różnic w geometrii wiązania przyjętych przez badane 8-aza-puryny. Ligandy te mogą eksponować różne atomy azotu (nr 7, 8 lub 9) pierścienia triazolowego na kanał prowadzący do miejsca wiązania, co może powodować rybozylację ligandu w tych eksponowanych pozycjach. Złożoność tego procesu potęguje obecność różnych tautomerycznych form ligandów, a także równowaga stanów protonacyjnych reszt aminokwasowych obecnych w miejscu wiązania.

W tej rozprawie zaprezentowane zostały wyniki obliczeń metod chemii kwantowej zastosowanych do różnych form tautomerycznych z8izoGua oraz azaGua, co pozwoliło na przedstawienie obrazu równowagi tautomerycznej tej sondy fluorescencyjnej. Wyniki obliczeń zostały porównane z wynikami otrzymanymi dla Guaniny i izoGua. Metoda zastosowana dla z8izoGua oraz azaGua została także przetestowana na 2,6 – diamino-8-azapurynie (DaaPur). Poprawność naszej metodologii została sprawdzona przez obliczenie elektronowych stanów wzbudzonych oraz porównanie ich z dostępnymi danymi eksperymentalnymi [14,30,31]. Także wpływ metylacji na wybraną azapurynę –

z8izoGua – poprzez metylację pozycji N9 (9m-z8izoGua) został zbadany. Grupa metylowa imituje pierścień deoxyrybozy, zatem badany układ naśladuje molekułę z8izoGua umieszczoną w łańcuchu polinukleotydowym. 9m-z8izoGua ma o 12 więcej możliwych tautomerów porównując do metylowanej izoGua, co jest spowodowane większą ilością stanów protonacyjnych pierścienia triazolowego, co z kolei może wpływać na rozkład gęstości elektronowej badanej molekuły. Z tego powodu została przeprowadzona także analiza rozkładu ładunków Hirshfelda. Obliczenia zostały wykonane na tym samym poziomie teorii kwantowej jak ten zastosowany do z8izoGua [32] i porównane do wyników otrzymanych za pomocą metod kompozytowych Gaussian-3 [33] oraz Gaussian-4 [34].

W pracy podjęto również próbę wyjaśnienia mechanizmu kanonicznej i niekanonicznej rybozylacji dwóch azapuryn przez cielęce PNP, odpowiednio azaGua oraz DaaPur, który został zaobserwowany podczas badań eksperymentalnych [29]. Na podstawie ustalonej równowagi tautomerycznej tych dwóch związków wyselekcjonowano najbardziej prawdopodobne formy, które następnie zostały poddane procedurze dokowania w kieszeni wiążącej odpowiednio wybranej i przygotowanej konformacji białka. W wyniku dokowania odpowiednie mody wiązania zostały zakwalifikowane do dalszych badań. Następnie wykorzystano odpowiednie techniki modelowania molekularnego, aby odpowiedzieć na pytanie, dlaczego dwa badane analogi puryn są rybozylowane w różnych pozycjach. Wyniki zostały omówione w kontekście dostępnych danych eksperymentalnych.

Cel pracy doktorskiej

Celem pracy jest zbadanie właściwości wybranych 8-aza-puryn: 8-aza-izo-guaniny, 8-aza-guaniny, 2,6-diamino-azapuryny oraz wyjaśnienie mechanizmu niekanonicznej rybozylacji wybranych ligandów – pochodnych 8-aza-puryn – przez fosforylaze nukleozydów purynowych (PNP).

Omówienie prac wchodzących w skład rozprawy doktorskiej

Na niniejsza rozprawę doktorską składają się trzy oryginalne publikacje, które są spójnym tematycznie cyklem artykułów. Prace te zostały opublikowane w indeksowanych czasopismach naukowych o zasięgu międzynarodowym, znajdujących się w bazie Journal Citation Report i wypełniają założony plan projektu badawczego.

Artykułem otwierającym projekt badawczy jest praca pt. „Theoretical study of tautomeric equilibria of 2,6-diamino-8-azapurine and 8-aza-iso-Guanine”. W pracy wyznaczono obraz tautomeryczny dwóch wybranych pochodnych 8-azapuryń - 2,6-diamino-8-azapuryny (DaaPur) oraz 8-aza-izo-guaniny (z8izoGua) - w otoczeniu gazowym oraz wodnym. Badania zostały przeprowadzone za pomocą metod chemii kwantowej na poziomie B3LYP/6-311+G(d,p) oraz BHandHLYP/cc-pVTZ. Wszystkie badane molekuły to formy aminowe, gdyż jak wykazano w przypadku izoguaniny (izoGua), formy iminowe są mocno niekorzystne z energetycznego punktu widzenia [16]. Procedura wyznaczenia równowagi tautomerycznej badanych związków zarówno w gazie jak i w wodzie była następująca:

- W pierwszym kroku geometria pięciu form tautomerycznych DaaPur oraz jedenastu tautomerów azaGua w gazie została zoptymalizowana. W celu potwierdzenia znalezienia się cząsteczek w minimum energetycznym została przeprowadzona analiza wibracyjna. Geometrie tautomerów zoptymalizowanych w próżni zostały użyte jako startowe struktury w środowisku wodnym
- Następnie wykonano obliczenia termochemii, które pozwoliły wyznaczyć parametry termodynamiczne tautomerów, tj. energię elektronową, entalpię, energię swobodną Gibbsa, wkład entropowy oraz energię swobodną solwatacji. Analiza populacyjna została przeprowadzona na podstawie rozkładu Boltzmann
- W celu walidacji rozważań teoretycznych pierwsze trzy singletowe energie wzbudzenia zostały wyznaczone za pomocą czasowo-zależnej metody DFT oraz porównane z eksperymentalnie wyznaczonymi maksimami absorpcji. Wszystkie obliczenia stanów wzbudzonych zostały wykonane dla zoptymalizowanych geometrii stanu podstawowego

Wyniki analizy populacyjnej w gazie dla DaaPur wskazały na dominację tautomeru D9, który miał wyraźną przewagę nad kolejną formą – D8 ($\Delta G = 5.2$ kcal/mol). Populacja pozostałych tautomerów wydają się być mocno związana z odległością

między tautomerycznymi protonami a grupami aminowymi, na co wskazuje bardzo duża wartość ΔG dla formy D1, w której proton jest zlokalizowany pomiędzy dwoma grupami aminowymi. Solwatacja DaaPur znacząco obniża różnicę energii swobodnej Gibbsa pomiędzy tautomerami zajmującymi skrajne miejsca w kolejności energetycznej (ΔG spadła z ok. 28 do ok. 12 kcal/mol), jednak D9 nadal dominuje nad pozostałymi formami ($\Delta G = 3.9$ kcal/mol). Wyniki obliczeń stanów wzbudzonych wskazują na bardzo dobrą zgodność przeskalowanych stanów singletowych S1 oraz S3 tautomeru większościowego D9 z maksimami absorpcji otrzymanymi eksperymentalnie.

Analiza energii swobodnych Gibbsa dla z8izoGua w środowisku gazowym sugeruje niewielką przewagę form enolowych A29c oraz A29t ($\Delta G < 1$ kcal/mol) nad amino-tautomerem A38. Dominujące tautomery enolowe są bardziej stabilne niż formy oxo – A39 oraz A19. Z biologicznego punktu widzenia najważniejszą podgrupą tautomerów są te protonowane w pozycji N(9)-H, które mogą służyć jako modelowe zasady związane z pierścieniem cukru w łańcuchu kwasu nukleinowego. Ta ważna podgrupa tautomerów została zdominowana przez formy enolowe. Poza pierwszymi trzema wyraźnie dominującymi tautomerami, obecność pozostałych form w gazie jest niewielka bądź marginalna. Zanurzenie w środowisku wodnym azaGua powoduje, iż najbardziej stabilnym tautomerem staje się A38, chociaż solwatacja jest bardziej korzystna dla większości pozostałych tautomerów. Dla tautomerów dominujących w gazie – A29c, A29t – solwatacja okazała się szczególnie niekorzystną ($\Delta\Delta G > 5$ kcal/mol). Zyskują natomiast ważne biologicznie formy oxo – A39, A19. Ich obecność w środowisku wodnym jest znacząca, na co wskazuje $\Delta G < 3$ kcal, z niewielką przewagą A39 ($\Delta G = 0.6$ kcal/mol). Bliskość energetyczna A39 oraz A19 sprawia, iż ten drugi tautomer, który może formować interfejs Watsona-Cricka z izocytozyną, musi rywalizować z tautomerem N(3)-H. Różnica energii swobodnej Gibbsa pomiędzy najniżej i najwyżej energetycznym tautomerem zmalała z 22.3 kcal/mol do 11.0 kcal/mol. Pomimo najkorzystniejszej energii swobodnej solwatacji tautomeru A17 ($\Delta\Delta G = -10.5$ kcal/mol), wciąż jego populacja pozostaje bardzo niska. Analiza stanów wzbudzonych w wodzie wskazuje, iż wyznaczone teoretycznie przeskalowane energetyczne stany singletowe S1 oraz S3 formy A38 są bardzo zbliżone do tych otrzymanych eksperymentalnie. Jest to zgodne z wynikami obliczeń równowagi tautomerycznej, która wskazuje na tautomer A38 jako większościowy.

Zarówno w środowisku gazowym jak i wodnym wkłady entropowe dla obu związków nie przekroczyły 1.2 kcal/mol, co świadczy o izoentropowości procesu tautomeryzacji.

Drugim artykułem wchodzącym w skład pracy doktorskiej jest praca pt. „Theoretical investigations of tautomeric equilibrium of 9-methyl-8-azaiso-Guanine and its electrostatic properties”. Praca zawiera wyniki badań nad wpływem metylacji z8izoGua w pozycji N(9) na jego równowagę tautomeryczną a także nad jego właściwościami elektrostatycznymi. Obliczenia zostały przeprowadzone za pomocą metod wykorzystujących hybrydowe funkcjonały (BHandHLYP/cc-pVTZ) oraz metody kompozytowe (G3, G4). Wszystkie możliwe formy tautomeryczne 9m-z8izoGua zostały rozważone, w tym formy iminowe. Badania zarówno w środowisku gazowym jak i wodnym zostały przeprowadzone według poniższej procedury:

- Geometrie 34 tautomerów 9m-z8izoGua zostały zoptymalizowane. Następnie geometrie poszczególnych form zostały poddane procedurze weryfikacji minimum energetycznego poprzez analizę wibracyjną. Geometrią startową dla obliczeń w wodzie była zoptymalizowana struktura gazowa.
- W kolejnym kroku, na podstawie obliczeń termodynamicznych zostały wyznaczone parametry termodynamiczne tautomerów: energię elektronową, entalpię, energię swobodną Gibbsa, wkład entropowy oraz energię swobodną solwatacji. Rozkład populacji tautomerów 9m-izoGua został wyznaczony za pomocą rozkładu Boltzmanna.
- W celu zbadania właściwości elektrostatycznych badanego układu ładunki cząstkowe Hirshfelda oraz momenty dipolowe zostały wyznaczone.

Wyniki obliczeń równowagi tautomerycznej 9m-z8izoGua w fazie gazowej wskazują na dominację dwóch form amino-enolowych AEC(t). Forma cis jest nieco bardziej preferowana z energią swobodną niższą o ok. 1 kcal/mol. Być może powodem tej nieznacznej różnicy jest odległość pomiędzy grupą metylową a protonem O(2)-H. Populacja pozostałych tautomerów jest właściwie nieznaczająca. Ich energia swobodna (uwzględniając również biologicznie istotne tautomery AO1, AO3 oraz ItO13) jest co najmniej o 4 kcal/mol wyższa od najbardziej prawdopodobnego tautomeru. Warto zaznaczyć, że wśród trzech podgrup (amino-oxo, imino-oxo, imino-enol), niemalże wszystkie formy protonowane w pozycjach N(7) oraz N(8) są najbardziej dyskryminowane w ramach danej podgrupy. Tak jak w przypadku równowagi cis-trans, powodem tego zjawiska może być odległość pomiędzy protonami N(7)-H oraz N(8)-H a grupą metylową. Umieszczenie 9m-aza-izoGua w środowisku wodnym znacząco zreorganizowało równowagę tautomeryczną ujawnioną w gazie. Najbardziej preferowanymi tautomerami są amino-oxo formy protonowane w pozycji N(1) oraz N(3) (AO1 oraz AO3) z bardzo niewielką przewagą tego pierwszego. Jako że izoGua może

formować interfejs Watsona-Cricka z izo-cytozyną (izoCyt), a także biorąc pod uwagę niewielką różnicę pomiędzy dwoma najniższymi energetycznymi tautomerami, można zakładać że forma protonowana w N(1), która może utworzyć stabilną parę typu Watson-Crick z izoCyt, musi rywalizować o proton z tautomerem N(3). Najbardziej prawdopodobne formy w środowisku gazowym to AEt(c), pomimo niekorzystnej energii solwatacji oraz niskich wartości momentu dipolowego ($\Delta\Delta G = 6.91$ kcal/mol oraz 8.02 kcal/mol; $p = 0.90$ D oraz 2.85 D, odpowiednio) ich obecność w środowisku wodnym jest nadal znacząca ($\Delta G \approx 1$ kcal/mol). Szczególnie obecność tautomeru AEc jest godna uwagi, ponieważ może ona oddziaływać z tyminą (Tym) poprzez interfejs Watsona-Cricka. Większość imino-oxo tautomerów charakteryzują się niższą energią solwatacji niż forma AO1. Jak się spodziewano również formy imino-enolowe, pomimo relatywnie korzystnej solwatacji, nadal pozostają niestabilne. Izoentropowy charakter procesu tautomerizacji zarówno w środowisku gazowym jak i wodnym potwierdza stosunkowo niewielki wkład entropowy ($T\Delta S < 1$ kcal/mol). Elektryczny moment dipolowy najbardziej preferowanych form zarówno w fazie gazowej (AEc(t)) jak i w wodzie (AO1 oraz AO3) nie przekracza 8D. Ponadto suma ładunków cząstkowych na każdym z pierścieni dla tych tautomerów oscyluje wokół zera. Z drugiej strony wartość momentu dipolowego form protonowanych w pozycjach N(7) lub N(8) w większości wypadków jest wysoka (od ok. 7 D do ok. 24 D). Pierścienie tych specyficznych form są przedmiotem nierównomiernego rozkładu ładunku (od -1.07 a.u. na pierścieniu pirydyminowym IctO78 do 0.76 a.u. na pierścieniu triazolowym IctO78). Badany układ tautomeryczny zawiera kilka tautomerów, które mogą zostać zdefiniowane jako cwiterjony (AO7, IctEc(t)7, Ict(t)O17, Ict(t)O37, Ict(t)O38 oraz Ict(t)O78). Charakter cwiterjonowy tych molekuł może zostać rozpoznany po zwykłej analizie macierzy rzędów wiązań kowalencyjnych. Jednak są również takie formy tautomeryczne (AO8, Ict(t)Ec(t)8 and Ict(t)O18), które okazują się być cwiterjonami na podstawie analizy rozkładu ładunków cząstkowych, a nie w wyniku analizy macierzy rzędów wiązań kowalencyjnych.

Ostatnim artykułem składającym się na dysertację doktorską jest praca pt. „Why Purine Nucleoside Phosphorylase Ribosylates 2,6-diamino-8-azapurine in non-Canonical Positions? A Molecular Modeling Study”. Badania przeprowadzone w ramach tej pracy skupiają się na wyjaśnieniu procesu alternatywnej (niekanonicznej) rybozylacji – 2,6-diamino-8-azapuryny (DaaPur) w pozycji nr 7 lub 8 przez fosforylazę nukleozydów purynowych (PNP). W celu walidacji naszych wyników również standardowa (kanoniczna) rybozylacja 8-azaguaniny (azaGua) w pozycji nr 9 pierścienia triazolowego przez to samo białko została zbadana. Metodologia zastosowana do badanych układów może zostać podzielona na cztery kroki:

- Wyznaczenie równowagi tautomerycznej azaGua z wykorzystaniem metod chemii kwantowej na poziomie BHandHLYP/cc-pVTZ. Tautomeria drugiego liganda – DaaPur - została zbadana w pierwszej pracy również za pomocą wyżej wymienionej metody i bazy
- Wyznaczenie stanów protonacyjnych miareczkwalnych reszt białka za pomocą serwera sieciowego h++ w eksperymentalnym pH (6.6)
- Dokowanie najniżej energetycznych tautomerów do struktur krystalicznych białka oraz selekcja odpowiednich modów ze względu na mapę wiązań wodorowych w kieszeni wiążącej białka, a także orientacji geometrycznej ligandów względem kanału prowadzącego do tej kieszeni
- Zastosowanie bardziej wyrafinowanych metodologii (Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) w połączeniu z Normal Mode Analysis (NMA)) do wybranych modów wiązania otrzymanych w wyniku dokowania. Metody te pozwalają na wyznaczenie względnych energii swobodnych Gibbsa modów wiązania

W pierwszej pracy wchodzącej w skład pracy doktorskiej wyniki obliczeń równowagi tautomerycznej DaaPur wskazują na wyraźną dyskryminację formy z protonem zlokalizowanym w pozycji nr 1 ($\Delta G \approx 12$ kcal/mol). Wyniki zastosowania tej samej metodologii dla azaGua pokazały, że wszystkie trzy najniżej energetycznie tautomery w środowisku wodnym są protonowane w pozycji nr 1 i różnią się protonacją pierścienia triazolowego, który może być uprotonowany we wszystkich trzech dostępnych pozycjach: 9 (najniższa energia), 8 ($\Delta\Delta G = 2.9$ kcal/mol) oraz 7 ($\Delta\Delta G = 4.4$ kcal/mol). Tylko populacyjnie znaczące ($\Delta G < 5$ kcal/mol) tautomery zostały zakwalifikowane do dokowania, tj. D8 oraz D9 dla DaaPur i A17, A18 oraz A19 dla azaGua. Najbardziej prawdopodobne mody wiązania DaaPur wewnątrz kieszeni wiążącej PNP eksponują azot N8, co stoi częściowo w zgodzie z wynikami eksperymentu, w którym postuluje się równowagę pomiędzy produktami rybozylowanymi w pozycjach N7 oraz N8. Pomimo że najniżej energetyczny mod eksponujący azot N7 jest dopiero szósty w kolejności pod względem energii swobodnej, to jego obecność in vitro nie może zostać całkowicie wykluczona. Niska pozycja tego modu wiązania w rankingu jest spowodowana największym spadkiem entropii spośród wszystkich badanych modów wiązania. Jeśli rozważymy tylko wkład entropowy, to ten mod wiążący awansuje na czwartą pozycję, tracąc do podium jedynie 0.1 kcal/mol. W przypadku drugiego badanego liganda – azaGua – najbardziej prawdopodobny mod wiązania eksponuje pozycję nr 9, co jest zbieżne z wynikami eksperymentu. Chociaż następny w kolejności mod eksponuje pozycję nr 8, to następujące po nim trzy mody wiązania eksponują również pozycję nr 9. Analiza kombinacji

energii swobodnych Gibbsa oraz entalpii ligandów, białek oraz kompleksów wskazuje na ekspozycję N8 dla DaaPur oraz N8 lub N9 dla azaGua. Kolejny istotny wynik tej pracy jest związany z kluczową resztą aminokwasową obecną w kieszeni wiążącej – Glu201. Obydwa miareczkowania teoretyczne wykonane za pomocą serwera sieciowego h++ oraz symulacji dynamiki molekularnej w stałym pH formy apo białka w środowisku wodnym pokazały, że pK_a tego łańcucha bocznego jest znacząco wyższe niż to wyznaczone eksperymentalnie dla układu modelowego, a zatem populacja protonowanej formy nie jest zaniedbywalna i musi zostać wzięta pod uwagę w dokowaniu. Analiza populacji tautomerów azaGua oraz DaaPur w połączeniu z teoretycznym miareczkowaniem oraz analiza dokowanie/MM-PBSA/NMA pokazały, że analogi puryn z podstawnikami 6-oxo i 6-amino powinny wiązać się do białek znajdujących się w różnych stanach protonacyjnych kieszeni wiążącej. Nasze wyniki sugerują, że azaGua (6-oxo podstawnik) wiąże się do kieszeni ze zdeprotonowanym Glu-201 (forma ujemnie naładowana), natomiast DaaPur (6-amino podstawnik) wiąże się do kieszeni z uprotonowanym Glu-201 (forma neutralna), która zgodnie z naszymi teoretycznymi wynikami miareczkowania powinna być znacząco populowana w środowisku wodnym.

Wnioski

Tautomeria 8-aza-izo-guaniny (z8izoGua), 2,6-diamino-aza-puryny (DaaPur), 8-aza-guaniny (azaGua)

- Najbardziej stabilnym tautomerem DaaPur w wodzie jest forma protonowana w pozycji N(9)
- Amino-oxo tautomer azaGua protonowany w pozycjach N(1) oraz N(9) jest najbardziej stabilnym tautomerem w wodzie
- Spośród tautomerów z8izoGua, najniższą energetyczną formą A38 jest ta, której maksima absorpcji znajdują się najbliżej przeskalowanych energii wzbudzenia, silnie wskazując protonację pozycji N(3) pierścienia pirymidynowego, czego wyniki eksperymentalne nie wyjaśniają
- W porównaniu do guaniny kolejność energetyczna form A19 oraz A39 z8izoGua jest odwrócona, a więc efekt mutacji izoCyt→Tym zaobserwowany dla alternatywnego kodu genetycznego może być spodziewany również przypadku zamiany izoGua na z8izoGua

Metylacja z8izoGua

- Dla metylowanej w pozycji N(9) z8izoGua formy amino-oxo AO1 oraz AO3 dominują w środowisku wodnym
- Metylacja z8izoGua znacząco zwiększa populację amino-enolowych form Aec(t). Ten efekt nie jest spowodowany solwatacją, lecz raczej zmianami w strukturze elektronowej molekuly
- Jako że tautomer AEc może być zarówno donorem jak i akceptorem wiązania wodorowego, może to skutkować utworzeniem trzech wiązań wodorowych z tyminą. Znacząca obecność AEc może doprowadzić do mutacji genetycznej izoCyt → Tym
- Tautomery 9m-izoGua z uprotonowanym pierścieniem triazolowym wykazują charakter cwiterjonowy
- Dla kilku tautomerów niemożliwe jest rozpoznanie cwiterjonowości na podstawie analizy rzędów wiązań. Ich cwiterjonowy charakter może być potwierdzony tylko poprzez analizę rozkładu ładunków wyznaczoną za pomocą metod chemii kwantowej

Powinowactwo DaaPur oraz azaGua do PNP

- Dwa najbardziej prawdopodobne mody wiązania DaaPur wewnątrz kieszeni wiążącej cieleńcego PNP eksponują azot N8, co jest częściowo zgodne z wynikami eksperymentu, które wskazują na równowagę pomiędzy produktami rybozylowanymi w pozycjach N7 i N8
- W przypadku azaGua najbardziej prawdopodobny mod wiązania eksponuje pozycję N9, co jest zbieżne z wynikami eksperymentu. Jak można zauważyć drugi w kolejności mod wiązania eksponuje pozycję N8, jednak następne trzy mody wiązania wskazują na potencjalną rybozylację w pozycji N9
- pK_a kluczowej reszty aminokwasowej w kieszeni wiążącej – Glu-201 – jest znacząco wyższe niż to wyznaczone eksperymentalnie dla układu modelowego, a zatem populacja protonowanej formy nie jest zaniedbywalna i musi zostać wzięta pod uwagę w procedurze dokowania
- azaGua (6-oxo podstawnik) wiąże się z białkiem zawierającym zdeprotonowaną resztę Glu-201, natomiast DaaPur (6-amino podstawnik) z białkiem zawierającym uprotonowaną resztę Glu-201
- Porównanie wyznaczonych modów wiązania badanych ligandów z inhibitorami obecnymi w strukturach krystalicznych sugeruje, że modyfikacja inhibitora (S)-PMPDAP, w którym łańcuch 2- (fosfonometoksy) propylowy jest przyłączony w pozycji 8 zamiast w pozycji 9, może zwiększyć jego powinowactwo wiązania się do białka PNP
- Podział modów wiązania ligandów na cztery rodziny ekspozycyjne (eksponujące atomy N3, N7, N8 i N9) oraz analiza ich zmian energii swobodnej Gibbsa w procesie wiązania do białka PNP wskazuje na ekspozycję N8 dla DaaPur oraz N8 lub N9 dla azaGua. Analiza rodzin ekspozycyjnych uwzględnia populacje różnych form tautomerycznych ligandów oraz form protonacyjnych reszty Glu-201

Literatura

- [1] J.D. Watson, F.H.C. Crick, *Nature* (1953), **171**: 737.
- [2] W. Seanger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, 1984.
- [3] J.S. Kwiatkowski, J. Leszczyński, *J. Mol. Struct.: THEOCHEM* (1990), **208**: 35.
- [4] M.T. Chenon, R.J. Pugmire, D.M. Grant, R.P. Panzica, L.B. Townsend, *J. Am. Chem. Soc.* (1975), **97**: 4636.
- [5] A.O. Alyoubi, R.H. Hilal, *Biophys. Chem.* (1995), **55**: 231.
- [6] J. Lin, C. Yu, S. Peng, I. Akiyama, K. Li, L.K. Lee, P.R. LeBreton, *J. Am. Chem. Soc.* (1980), **102**: 4627.
- [7] G. Fogarasi, *Chem. Phys.* (2008), **349**: 204.
- [8] D. Kosenkov, Y. Kholod, L. Gorb, O. Shishkin, D.M. Hovorun, M. Mons, J. Leszczyński, *J. Phys. Chem.* (2009), **113**: 6140.
- [9] M. Sabio, S. Topiol, W.C. Lumma, *J. Phys. Chem.* (1990), **94**: 1366.
- [10] H.-S. Kim, D.-S. Ahn, S.-Y. Chung, S.K. Kim, S. Lee, *J. Phys. Chem.* (2007), **111**: 8012.
- [11] C. Plutzer, K. Kleinermanns, *Phys. Chem. Chem. Phys.* (2002), **4**: 4877.
- [12] C. Roberts, R. Bandaru, C. Switzer, *J. Am. Chem. Soc.* (1997), **119**: 4640.
- [13] I. Hirao, *Curr. Opin. Chem. Biol.* (2006), **10**: 622.
- [14] J. Wierzchowski, G. Mędza, J. Sepioł, M. Szabelski, D. Shugar, *J. Photochem. Photobiol. A* (2012), **237**: 64.
- [15] J. Wierzchowski, J.M. Antosiewicz, D. Shugar, *R. Soc. Chem.* (2014), **10**: 2756.
- [16] J.R. Blas, F.J. Luque, M. Orozco, *J. Am. Chem. Soc.* (2004), **126**: 154.
- [17] A. Bzowska, E. Kulikowska, D. Shugar, *Pharmacol. Ther.* (2000), **88**: 349.
- [18] J.D. Stoeckler, *Dev. Cancer Chemother.* (1984), 35.
- [19] R.M. Mader, A. E. Sieder, J. Braun, B. Rizovski, M. Kalipciyan, M. W. Mueller, R. Jakesz, H. Rainer, G.G. Steger, *Biochem. Pharmacol.* (1997), **54**: 1233.
- [20] J.A. Secrist, W.B. Parker, P.W. Allan, L.L. Bennett, W.R. Waud, J.W. Truss, A. T. Fowler, J.A. Montgomery, S.E. Ealick, A.H. Wells, G.Y. Gillespie, V.K. Gadi, E.J. Sorscher, *Nucleotides Nucleosides* (1999), **18**: 745.
- [21] F. Ravandi, V. Gandhi, *Expert Opin. Investig. Drugs* (2006), **15**: 1601.
- [22] T. Robak, E. Lech-Maranda, A. Korycka, E. Robak, *Med. Chem.* (2009), **13**: 3165.
- [23] M.C. Ho, W. Shi, A. Rinaldo-Matthis, P.C. Tyler, G.B. Evans, K. Clinch, S.C. Almo, V.L. Schramm, *Proc. Natl. Acad. Sci. USA* (2010), **10**: 4805.
- [24] I.A. Mikhailopulo, *Curr. Org. Chem.* (2007), **11**: 317.
- [25] I.A. Mikhailopulo, A.I. Miroshnikov, *Mendeleev Commun.* (2011), **21**: 57.

- [26] V.A. Stepchenko, F. Seela, R.S. Esipov, A.I. Miroshnikov, Y.A. Sokolov, I.A. Mikhailopulo, *Synlett* (2012), **23**: 1541.
- [27] A. Stachelska-Wierzchowska, J. Wierzchowski, B. Wielgus-Kutrowska, G. Mikleušević, *Molecules* (2013), **18**: 12587.
- [28] C.P. Da Costa, M.J. Fedor, L.G. Scott, *J. Am. Chem. Soc.* (2007), **129**: 3426.
- [29] A. Stachelska-Wierzchowska, J. Wierzchowski, A. Bzowska, B. Wielgus-Kutrowska, *Molecules* (2016), **21**: 44.
- [30] J. Wierzchowski, G. Mędza, M. Szabelski, A. Stachelska-Wierzchowska, *J. Photochemis. Photobiol. A* (2013), **265**: 49.
- [31] G. Mędza, Ph.D. thesis (2014).
- [32] M. Pyrka, M. Maciejczyk, *Chem. Phys. Lett.* (2015), **627**: 30.
- [33] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rassolov, J.A. Pople, *J. Chem. Phys.* (1998), **109**: 7764.
- [34] L.A. Curtiss, P.C. Redfern, K. Raghavachari, *J. Chem. Phys.* (2007), **126**: 084108.

III. Streszczenie w języku polskim

Równowaga tautomeryczne form aminowych 2,6-diamino-8-azapuryny, 8-aza-izo-guaniny oraz 8-azaguaniny została wyznaczona za pomocą obliczeń DFT na poziomie B3LYP/6311+G(d,p) oraz BHandHLYP/cc-pVTZ. Najbardziej prawdopodobny tautomer 2,6-diamino-8-azapuryny jest protonowany w pozycji N(9), co jest zbieżne z wynikami dostępnymi danymi eksperymentalnymi. Tautomer o najniższej energii swobodnej 8-aza-izo-guaniny jest protonowany w pozycjach N(3) i N(8). W przypadku ważniejszego biologicznie tautomeru N(9)-H prawdopodobieństwo protonacji w pozycji N(3) jest wyższe niż w N(1). Ten wynik, obserwowany również dla izo-guaniny, pokazuje odwrócone prawdopodobieństwo protonacji w pozycjach N(3) i N(1) w porównaniu z guaniną, a więc efekt mutacji izocytozyna->tymina, zaobserwowany dla alternatywnego kodu genetycznego, może również mieć miejsce w przypadku, kiedy izo-guanina zostanie zastąpiona 8-aza-izo-guaniną. Wyraźnie dominującym tautomerem 8-aza-guaniny jest forma uprotonowana w pozycji N(1) oraz N(9), co jest zgodne z wynikami otrzymanymi dla guaniny. Zatem guanina może z powodzeniem zostać zastąpiona przez 8-aza-guaninę, ponieważ nie niesie to ze sobą konsekwencji w postaci zaburzenia interfejsu Watsona-Cricka. Obliczone energie stanów wzbudzonych dla 8-aza-izo-guaniny oraz 2,6-diamino-8-azapuryny pozostają w dobrej zgodności z dostępnymi danymi eksperymentalnymi i wskazują na najbardziej prawdopodobne tautomery, jako te odpowiedzialne za absorpcję.

Obraz tautomeryczny 9-metylo-8-aza-izo-guaniny został zbadany również metodami hybrydowej chemii kwantowej (BHandHLYP/cc-pVTZ) oraz dodatkowo metodami kompozytowymi (G3, G4). Najbardziej dominującymi tautomerami są formy amino-oxo, które są protonowane w pozycjach N(1) i N(3). Ta kolejność dominacji jest zgodna z wynikami tautomerii otrzymanych dla 8-aza-izo-guaniny, ale metylacja badanej cząsteczki powoduje znaczny wzrost populacji form amino-enolowych protonowanych w pozycji O(2). Analiza właściwości elektrostatycznych 9-metylo-8-aza-izo-guaniny pokazuje, że dominujące tautomery mają równomierny rozkład ładunku, ale zdecydowana większość wysokoenergetycznych form protonowanych na pierścieniu triazolowym to cwiterjony.

Fosforylaza nukleozydów purynowych (PNP) jest enzymem, który katalizuje odwracalny proces konwersji (rybozylacja i fosforoliza) między zasadami nukleinowymi (purynami) i ich nukleozydami. Badania eksperymentalne

wykazały, że cielęce PNP rybozyluje analogi purynowe w określonych pozycjach - 2,6-diamino-8-azapuryna w pozycjach 7 lub 8 i 8-azaguanina w pozycji 9 pierścienia triazolowego. Przyczyną tego zjawiska może być różna ekspozycja substratów purynowych na kanał prowadzący do miejsca wiązania. Tę hipotezę zweryfikowano przez zastosowanie technik modelowania molekularnego do dwóch kompleksów analogów purynowych 2,6-diamino-azapuryny - cielęce PNP (kod pdb: 1LVU) i 8-azaguaniny - cielęce PNP (kod pdb: 2A11). Wyniki uzyskane w połączeniu z metodami chemii kwantowej, dokowania i dynamiki molekularnej wykazały jakościową trafność naszej hipotezy. Energie swobodne wiązania układów białko-ligand pokazały, że najbardziej prawdopodobne mody wiązania eksponują azot N(8) w przypadku kompleksu z 2,6-diamino-8-azapuryną i azot N9 dla kompleksu z 8-azaguaniną w kierunku kanału wiążącego, a także wykluczają ekspozycję N9 dla 2,6-diamino-8-azapuryny i N7 dla 8-azaguaniny, co jest częściowo zgodne z danymi eksperymentalnymi. Innym ważnym wynikiem uzyskanym w tym badaniu jest znacznie większa populacja protonowanej formy kluczowej reszty Glu-201 (ze względu na jej obecność w kieszeni wiążącej), w porównaniu ze standardową protonacją niezwiązanego kwasu glutaminowego w roztworze. Wynik ten w połączeniu z populacjami postaci tautomerycznych obu badanych układów silnie sugeruje, że 2,6-diamino-8-azaguanina i 8-azaguanina są rozpoznawane przez białka odpowiednio ze zdeprotonowaną i protonowaną resztą Glu-201. Porównanie wyznaczonych modów wiązania badanych ligandów z inhibitorami obecnymi w strukturach krystalicznych sugeruje, że modyfikacja inhibitora (S)-PMPDAP, w którym łańcuch 2- (fosfonometoksy) propylowy jest przyłączony w pozycji 8 zamiast w pozycji 9, może zwiększyć jego powinowactwo wiązania się do białka PNP.

IV. Streszczenie w języku angielskim

Tautomeric equilibrium of the 2,6-diamino-8-azapurine, 8-aza-iso-guanine and 8-azaguanine amino forms was determined by DFT calculations at B3LYP/6311+G(d,p) and BHandHLYP/cc-pVTZ level of theory. The most populated tautomer of 2,6-diamino-8-azapurine is protonated at position N(9), which is consistent with the results available from experimental data. The tautomer with the lowest free energy of 8-aza-iso-guanine is protonated at positions N(3) and N(8). In case of the biologically important N(9)-H tautomer, the probability of protonation at the N(3) position is higher than at N(1). This result, also observed for iso-guanine, shows the inverse probability of protonation at positions N(3) and N(1) compared to guanine, and thus the effect of the isocytosine->Thymine mutation, observed for the alternative genetic code, can also occur if iso-guanine is replaced by 8-aza-iso-guanine. The clearly dominant tautomer of 8-aza-guanine is the form protonated at the positions N(1) and N(9), which is consistent with the results obtained for guanine. Therefore, guanine can be successfully replaced by 8-aza-guanine, because it does not affect a Watson-Crick interface. The calculated energies of the excited states for 8-aza-iso-guanine and 2,6-diamino-8-azapurine are in good agreement with available experimental data and indicate the most likely tautomers as those responsible for absorption.

The tautomeric picture of 9-methyl-8-aza-iso-guanine was also examined by hybrid quantum chemistry (BHandHLYP /cc-pVTZ) and additionally by composite methods (G3, G4). The most dominant tautomers are amino-oxo forms that are protonated at positions N(1) and N(3). This order of dominance is consistent with the results of tautomerism obtained for 8-aza-iso-guanine, but the methylation of the molecule under investigation causes a significant increase in the population of amino-enol forms protonated at position O(2). Analysis of the electrostatic properties of 9-methyl-8-aza-iso-guanine shows that the dominant tautomers have uniform distribution of charge, but the vast majority of high-energy forms protonated on the triazole ring are zwitterions.

Purine nucleoside phosphorylase (PNP) is an enzyme, that catalyzes reversible conversion process (ribosylation and phosphorolysis) between nucleobases (purines) and their nucleosides. Experimental studies showed that *calf* PNP ribosylates purine analogs in specific positions – 2,6-diamino-8-azapurine in positions 7 or 8 and 8-azaguanine in position 9 of the triazole ring. The reason of this phenomena can be a result of different exposition of purine substrates to the channel leading to the binding site. This hypothesis was verified by application of molecular modelling techniques to two complexes of purine

analogs 2,6-diamino-azapurine – *calf* PNP and 8-azaguanine – *calf* PNP (pdb-code: 1LVU and 2AI1, respectively). The results obtained with combination of quantum chemistry, docking and molecular dynamics methods showed qualitative validity of our hypothesis. Binding free energies of protein-ligand systems showed that most probable binding poses expose N8 nitrogen for 2,6-diamino-8-azapurine and N9 nitrogen for 8-azaguanine into the binding channel and ruled out exposition of N9 for 2,6-diamino-8-azapurine and N7 for 8-azaguanine, partially in agreement with the experimental data. The other important result obtained in this study is a significantly higher population of protonated form of crucial residue Glu-201 present in the binding pocket, compared to the standard protonation of free glutamic acid in solution. This result combined with populations of tautomeric forms of both investigated systems strongly suggests that 2,6-diamino-8-azaguanine and 8-azaguanine is recognized by proteins with deprotonated and protonated Glu-201 residue, respectively. Comparison of computed binding poses of the investigated ligands to the inhibitors present in crystal structures suggests that modification of (S)-PMPDAP inhibitor, in which 2-(phosphonomethoxy) propyl chain is attached at position 8 instead of position 9, might increase its binding affinity.

V. Dorobek naukowy

Publikacje

Publikacje wchodzące w skład rozprawy doktorskiej:

Maciej Pyrka, Maciej Maciejczyk*, Theoretical study of tautomeric equilibria of 2,6-diamino-8-azapurine and 8-aza-iso-Guanine, *Chemical Physics Letters* (2015), **627**: 30.

Praca oryginalna

Punkty MNiSW: **25** (2015), **70** (2019)

Impact Factor: **1,897**

Maciej Pyrka, Maciej Maciejczyk*, Theoretical investigations of tautomeric equilibrium of 9-methyl-8-aza-iso-Guanine and its electrostatic properties, *Computational and Theoretical Chemistry* (2016), **1091**: 1-7.

Praca oryginalna

Punkty MNiSW: **20** (2016), **40** (2019)

Impact Factor: **1,403** (2016),

Maciej Pyrka, Maciej Maciejczyk*, Why purine nucleoside phosphorylase ribosylates 2,6-diamino-8-azapurine in non-canonical positions? A molecular modeling study, *Journal of Chemical Information and Modeling*, (2020), **1091**: 1.

Praca oryginalna

Punkty MNiSW: **100** (2019)

Impact Factor: **3,966** (2018)

Pozostałe publikacje:

F. Ferrari, **M. Pyrka**; Dynamical aspects of inextensible chains, *International Journal of Modern Physics B* (2012), **26**: 1250009

Praca oryginalna

Punkty MNiSW: **15** (2012), **20** (2019)

Impact Factor: **0,402** (2012)

B. Smyk, G. Mędrza, A. Kasperek, **M. Pyrka**, I. Gryczynski, M. Maciejczyk, Spectroscopic Properties and Conformational Analysis of Methyl Ester of Sinapic Acid in Various Environments, *Journal of Physical Chemistry B* (2017), **121**: 7299

Praca oryginalna

Punkty MNiSW: **30** (2017) **100** (2019)

Impact Factor: **3,146** (2017)

A. Krężel, M. Padjasek, M. Maciejczyk, M. Nowakowski, O. Kerber, **M. Pyrka**, W. Koźmiński, Metal exchange in interprotein Zn(II) binding site of hook domain – structural insights into Cd(II)-induced DNA repair inhibition, *Chemistry – A European Journal*, (2020).

Praca oryginalna

Punkty MNiSW: **140** (2019)

Impact Factor: **5,160** (2018)

Sumaryczny IF prac: **15,974** (470 pkt MNiSW)

Działalność naukowa i organizacyjna

Udział w konferencjach:

- Krajowa Konferencja Zastosowań Matematyki w Biologii i Medycynie Zakopane, M. Pyrka, P. Boguś, Dyskretne i nierozciągliwe modele łańcucha polimerowego, 2011, poster
- XVI Zjazd Polskiego Towarzystwa Biofizycznego, konferencja naukowa, Ryn, M.Pyrka, M. Maciejczyk, Theoretical study of tautomeric equilibria of 8-azapurine amino forms, 2016, poster

- XVII Zjazd Polskiego Towarzystwa Biofizycznego, konferencja naukowa, Olsztyn, M. Pyrka, M. Maciejczyk Theoretical investigations of alternative ribosylation process of selected 8-azapurines by purine nucleoside phosphorylase, 2019, prezentacja

Udział w warsztatach:

- Warsztaty pt. „Cząsteczka w otoczeniu chemicznym”, Wrocław, 2015

- Warsztaty dotyczące oprogramowania Gaussian służącemu przeprowadzeniu obliczeń kwantowych na niewielkich cząsteczkach chemicznych, Santiago de Compostela, Hiszpania, 2017

Staże:

- Miesięczny staż naukowy w Zakładzie Teorii Pola, Instytut Fizyki, Uniwersytet Szczeciński, 2010, Szczecin

- Trzymiesięczny staż naukowy w “The Roitberg Lab” na University of Florida, 2018, Gainesville, USA

Udział w komitetach organizacyjnych:

- Komitet Organizacyjny XVI oraz XVII Zjazdu Polskiego Towarzystwa Biofizycznego, 2014-2019

Nagrody:

- Nagroda zespołowa II stopnia Rektora Uniwersytetu Warmińsko – Mazurskiego w Olsztynie za osiągnięcia w dziedzinie dydaktycznej

- Nagroda za najlepszą prezentację podczas XVII Zjazdu Polskiego Towarzystwa Biofizycznego, Olsztyn, 2019

VI. Oświadczenia współautorów prac wchodzących w skład rozprawy doktorskiej

Mgr Maciej Pyrka
Katedra Fizyki i Biofizyki
Uniwersytet Warmińsko-Mazurski

Oświadczenie współautora

Dotyczy publikacji:

Maciej Pyrka, Maciej Maciejczyk, Theoretical study of tautomeric equilibria of 2,6-diamino-8-azapurine and 8-aza-iso-Guanine, *Chemical Physics Letters* (2015), **627**: 30.

Oświadczam, że mój wkład w powstanie publikacji polegał na współudziale przy tworzeniu pracy, zebraniu, analizie i interpretacji zgromadzonej literatury oraz przygotowaniu tekstu manuskryptu.

Mój udział oceniam na 70%

podpis


.....

Mgr Maciej Pyrka
Katedra Fizyki i Biofizyki
Uniwersytet Warmińsko-Mazurski

Oświadczenie współautora

Dotyczy publikacji:

Maciej Pyrka, Maciej Maciejczyk, Theoretical investigations of tautomeric equilibrium of 9-methyl-8-aza-iso-Guanine and its electrostatic properties, *Computational and Theoretical Chemistry* (2016), **1091**: 1.

Oświadczam, że mój wkład w powstanie publikacji polegał na współudziale przy tworzeniu pracy, zebraniu, analizie i interpretacji zgromadzonej literatury oraz przygotowaniu tekstu manuskryptu.

Mój udział oceniam na 70%

podpis


.....

Mgr Maciej Pyrka
Katedra Fizyki i Biofizyki
Uniwersytet Warmińsko-Mazurski

Oświadczenie współautora

Dotyczy publikacji:

Maciej Pyrka, Maciej Maciejczyk, Why Purine Nucleoside Phosphorylase Ribosylates 2,6-Diamino-8-azapurine in Noncanonical Positions? A Molecular Modeling Study, *Journal of Chemical Information and Modeling* (2020), **60**: 1595.

Oświadczam, że mój wkład w powstanie publikacji polegał na współudziale przy tworzeniu pracy, zebraniu, analizie i interpretacji zgromadzonej literatury oraz przygotowaniu tekstu manuskryptu.

Mój udział oceniam na 80%

podpis


.....

Dr hab. Maciej Maciejczyk
Katedra Fizyki i Biofizyki
Uniwersytet Warmińsko-Mazurski

Oświadczenie współautora

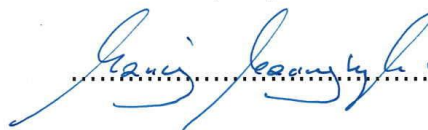
Dotyczy publikacji:

Maciej Pyrka, **Maciej Maciejczyk**, Theoretical study of tautomeric equilibria of 2,6-diamino-8-azapurine and 8-aza-iso-Guanine, *Chemical Physics Letters* (2015), **627**: 30.

Oświadczam, że mój wkład w powstanie publikacji polegał na współdziałaniu przy tworzeniu koncepcji pracy, pomocy merytorycznej przy tworzeniu pracy oraz nadaniu pracy ostatecznej formy.

Mój udział oceniam na 30%

podpis



Dr hab. Maciej Maciejczyk
Katedra Fizyki i Biofizyki
Uniwersytet Warmińsko-Mazurski

Oświadczenie współautora

Dotyczy publikacji:

Maciej Pyrka, **Maciej Maciejczyk**, Theoretical investigations of tautomeric equilibrium of 9-methyl-8-aza-iso-Guanine and its electrostatic properties, *Computational and Theoretical Chemistry* (2016), **1091**: 1.

Oświadczam, że mój wkład w powstanie publikacji polegał na współdziałaniu przy tworzeniu koncepcji pracy, pomocy merytorycznej przy tworzeniu pracy oraz nadaniu pracy ostatecznej formy.

Mój udział oceniam na 30%

podpis



Dr hab. Maciej Maciejczyk
Katedra Fizyki i Biofizyki
Uniwersytet Warmińsko-Mazurski

Oświadczenie współautora

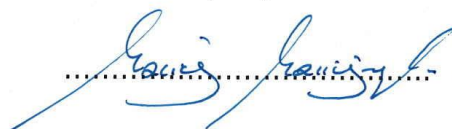
Dotyczy publikacji:

Maciej Pyrka, **Maciej Maciejczyk**, Why Purine Nucleoside Phosphorylase Ribosylates 2,6-Diamino-8-azapurine in Noncanonical Positions? A Molecular Modeling Study, *Journal of Chemical Information and Modeling* (2020), **60**: 1595.

Oświadczam, że mój wkład w powstanie publikacji polegał na współudziale przy tworzeniu koncepcji pracy, pomocy merytorycznej przy tworzeniu pracy oraz nadaniu pracy ostatecznej formy.

Mój udział oceniam na 20%

podpis



VII. Kopie publikacji składających się na rozprawę doktorską

Chemical Physics Letters 627 (2015) 30–35



Contents lists available at ScienceDirect

Chemical Physics Letters

journal homepage: www.elsevier.com/locate/cplett



Theoretical study of tautomeric equilibria of 2,6-diamino-8-azapurine and 8-aza-iso-Guanine



Maciej Pyrka^a, Maciej Maciejczyk^{a,b,*}

^a Department of Physics and Biophysics, Faculty of Food Science, University of Warmia and Mazury, ul. Oczapowskiego 4, 10-719 Olsztyn, Poland

^b Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, ul. Ks. Trojdena 4, 02-109 Warsaw, Poland

ARTICLE INFO

Article history:

Received 30 October 2014

In final form 17 March 2015

Available online 26 March 2015

ABSTRACT

Tautomeric equilibria of amino forms of 2,6-diamino-8-azapurine and 8-aza-iso-Guanine is revealed by DFT computations. The most populated tautomer of 2,6-diamino-8-azapurine, in agreement with available experimental data, is protonated at position N(9). The lowest free energy tautomer of 8-aza-iso-Guanine is protonated at positions N(3) and N(8). For biologically more important tautomer N(9)-H, probability of N(3) protonation is higher than N(1). This result, observed also for iso-Guanine, shows reversed probability of protonation at positions N(3) and N(1), compared to Guanine and decreased specificity of pairing of 8-aza-iso-Guanine with iso-Cytosine should be expected. Computed electronic excitation energies well match available experimental data.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The tautomeric equilibrium of nucleobases is one of key factors responsible for correct base–base recognition in nucleic acids [1] and its disturbance has a large impact on biological processes as it can lead to genome mutations [2]. The tautomerization process of purines and nucleobases has been extensively studied by both theoretical and experimental methods [3–11]. Development of the expanded genetic code raised natural question about tautomeric equilibrium of its new letters – iso-Guanine and iso-Cytosine [12].

Recently synthesized derivative of iso-Guanine (isoGua) called aza-iso-Guanine (z8isoGua) was recognized as a promising fluorescent probe [13], which can be used for studies of important biological processes. The replacement of C8 carbon of isoGua does not directly influence both Watson–Crick and Hoogsteen interfaces, although it can possibly change tautomeric equilibrium of the new compound compared to unmodified isoGua.

In this letter results of quantum chemistry methods applied to various tautomeric forms of z8isoGua are presented. The picture of tautomeric equilibrium of this fluorescent probe is revealed. The results are compared to those obtained for natural Guanine and isoGua. The method applied to z8isoGua was also tested on

2,6-diamino-8-azapurine (DaaPur) molecule. The reliability of our methodology was checked by computation of electronic excitation energies and their comparison with available experimental data [13–15].

2. Methods

Quantum chemistry methods have been successfully applied to the description of tautomerization phenomena in small molecules [16–18]. It was shown previously [16,19–23], that density functional theory (DFT) with B3LYP functional and 6-311+G (d,p) basis set [24–27] is sufficient for correct description of tautomerization phenomena and results generated by DFT method are close to those obtained with the higher level methods like MP2, MP4 and CCSD. Although chemical compounds studied in publications above do not include keto–enol tautomerism, characteristic for z8isoGua molecule. Therefore, only preliminary calculations in both the gas phase and water (PCM method [28,29]) were performed at the DFT/B3LYP method and 6-311+G (d,p) basis set level of theory with GAMESS package [30,31]. These calculations served as the preliminary sieve and structures of the lowest energy tautomers were also optimized with higher level DFT-BHandHLYP method [32] combined with cc-pVTZ basis set [33], the method/basis-set combination, which was successfully applied to benzoderivatives of nucleic acid bases [34]. Geometries of the selected tautomers were optimized in neutral form and ground state and the vibrational analysis of each tautomer was carried out. The latter step

* Corresponding author at: Department of Physics and Biophysics, Faculty of Food Science, University of Warmia and Mazury, ul. Oczapowskiego 4, 10-719 Olsztyn, Poland.

E-mail address: maciej.maciejczyk@uwm.edu.pl (M. Maciejczyk).

<http://dx.doi.org/10.1016/j.cplett.2015.03.029>

0009-2614/© 2015 Elsevier B.V. All rights reserved.

served two purposes. First, vibrational analysis proved that all structures reached energy minima in the optimization process. Second, the output of vibrational analysis contains information about thermochemistry of the investigated system. The optimization of water-solvated systems was speed up by using final structures from the optimization in vacuum as initial structures for optimization in solvent.

The following thermodynamic parameters were obtained from thermochemistry calculations: energy (E – including electronic energy and zero point energy), enthalpy ($H = E + k_B T$), entropy (S) and Gibbs free energy ($G = H - TS$). The tautomerization preference was determined by calculation of the relative thermodynamic parameters (ΔE , ΔH , ΔG , $T\Delta S$) with respect to the lowest-energy tautomer. The relative free energy ΔG depends on electronic energy, zero point energy, thermal corrections to energy and entropy. The solvation free energy $\Delta\Delta G$ was calculated according to the equation:

$$\Delta\Delta G = \Delta G_{\text{water}} - \Delta G_{\text{gas}} \quad (1)$$

where ΔG_{water} and ΔG_{gas} are relative free energies of selected tautomer in water and gas phase, respectively. Furthermore, tautomeric equilibrium constants ($pK = \Delta G/2.303RT$) and percentage contents of individual tautomer ($x = K/(1 + K)$) were computed.

The validity of our theoretical prediction was checked by comparison of the energies of electronic excitation with available positions of absorption peaks. For this purpose first three singlet electronic excited states (vertical states) were computed using time-dependent DFT [35] and BHandHLYP/cc-pVTZ functional/basis-set methodology. All excited-state calculations were performed for ground-state-optimized geometries. The computed excitation energies were scaled by factors specific for each tautomer in the same manner as described by Shukla et al. [36–38]. Scaling factors were set to match excitation energies of the experimentally determined and computed first singlet excited states (S1). Then agreement between scaled excitation energies of third singlet excited state (S3) and experimentally determined one was checked.

Various tautomers of DaaPur and z8isoGua molecules considered in this publication are shown in Figures 1 and 2, respectively. In both cases only tautomers with amino groups were considered. It was shown that tautomers with amino groups are dominant for both adenine [16] and isoGua [17], so imino tautomers are presumably less important for DaaPur and z8isoGua molecules, also. Nevertheless, calculations of imino and charged tautomers are in progress in our laboratory.

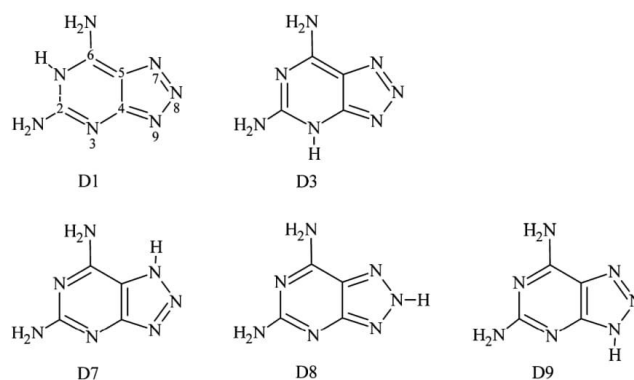


Fig. 1. DaaPur tautomeric structure. In contrast to z8isoGua, only one tautomeric proton is located on the entire structure.

Table 1

Relative thermodynamic parameters (kcal/mol) for DaaPur tautomers in the gas phase. For the assignment of symbols see Figure 1. The relative energies shown in parenthesis were computed with BHandHLYP/pVTZ level of theory.

Tautomer	ΔE	ΔH	ΔG	$T\Delta S$
D1	28.26	28.26	28.49	-0.24
D3	15.13	15.13	15.40	-0.55
D7	10.02	10.02	10.47	-0.45
D8	3.86 (5.13)	3.86 (5.13)	4.04 (5.16)	-0.18 (-0.18)

Positive values are related to less preferred tautomers than reference D9 tautomer. Thermochemistry was calculated at 298.15 K.

3. Results and discussion

3.1. DaaPur tautomeric system

The neutral amino form of DaaPur molecule is the simpler of two investigated tautomeric systems, as it has only one tautomeric proton, which can occupy one of five available positions as shown in Figure 1. The relative thermodynamic parameters such as: energy (ΔE), enthalpy (ΔH), free energy (ΔG), parameter related to entropy ($T\Delta S$) and free energy of solvation ($\Delta\Delta G$) of DaaPur molecule in gas phase and in water are collected in Tables 1 and 2, respectively.

Gas phase. The D9 tautomer is the most stable one and its free energy is around 4 kcal/mol lower than D8 – the second lowest energy tautomer. The population of various tautomers, shown in Table S1, seems to be closely related to the distance between tautomeric proton and two amino groups of DaaPur molecule. This observations are supported especially by a very large ΔG value obtained for D1 tautomer, in which proton is located between two amino groups of DaaPur molecule. Higher-level calculations performed for two lowest energy tautomers (D8 and D9) do not change their positions as D9 tautomer is still clearly dominant (see Table 1) and the energy gap between two lowest-energy tautomers increased by 1 kcal/mol.

Solvent effect. Immersion of DaaPur molecule in water significantly decreases the free energy difference between lowest (D9) and highest (D1) free energy tautomers, which dropped from 28.49 to 11.84 kcal/mol. D3 and D7 tautomers remain in the middle of the energy scale, although, because of a very strong solvation effect of D3 ($\Delta\Delta G = -10.33$ kcal/mol), they exchanged positions. It should be stressed here that, in agreement with the experimental data [14], D9 remains the lowest energy tautomer in water. As in the case of vacuum calculations, application of the higher-level method left

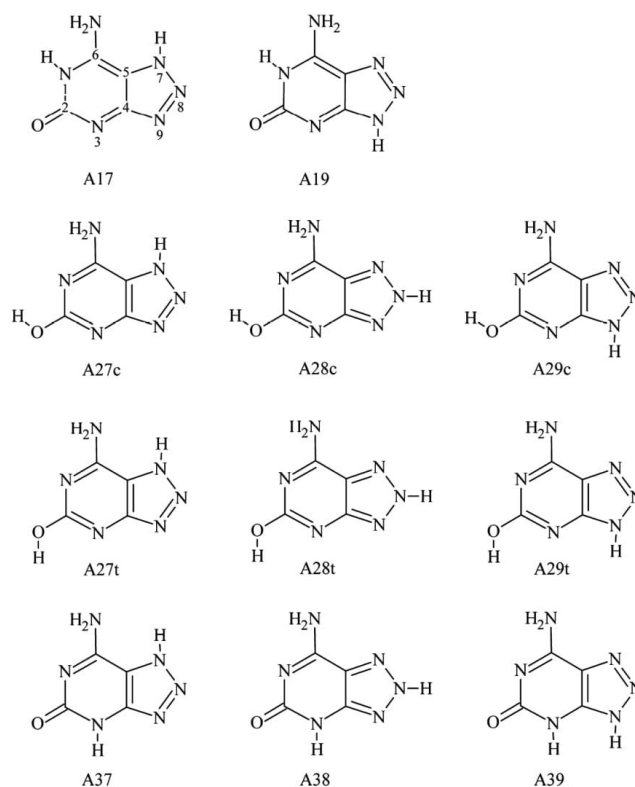


Fig. 2. z8isoGua tautomeric forms. First and last row present keto structures, while the second and third row show *cis* and *trans* enol structures, respectively.

the order of tautomers unchanged, and only increased the energy gap between two lowest-energy tautomers by 1 kcal/mol.

It was shown for many small molecules, that tautomeric conversions are isoentropic [16,19–23,39]. This observation can also be confirmed for DaaPur molecule in both gas phase and water, as all entropic terms $T\Delta S$ are smaller than 0.6 kcal/mol. Another specificity of the investigated system is, that all values of relative thermodynamic corrections to the energy, enthalpy and free energy are less than 1.0 kcal/mol for both vacuum and water (see Tables S1 and S2 in Supplementary Information), meaning that tautomeric conversions are rather weakly coupled to the temperature.

3.2. z8isoGua tautomeric system

The neutral amino form of z8isoGua molecule is the more complicated of two investigated tautomeric systems. Its two tautomeric

protons can exist in 11 possible forms as shown in Figure 2. The relative thermodynamic parameters such as: energy (ΔE), enthalpy (ΔH), free energy (ΔG), parameter related to entropy ($T\Delta S$) and free energy of solvation ($\Delta\Delta G$) of z8isoGua in gas phase and in water are collected in Tables 3 and 4, respectively.

Gas phase. The lowest energy tautomers, obtained with higher-level methodology, are A29c and A29t, closely followed by A38. Free energy differences between these forms are less than 1 kcal/mol. As the initial lower-level computations (B3LYP/6-311+G(d,p)) showed reversed order of tautomers, this result confirms necessity of application of the higher-level methods for keto-enol tautomeric systems. The A29c/t amino-enol tautomers are more stable than A39 and A19 amino-oxo forms as it was also predicted for isoGua [17]. The A29c tautomer is slightly more preferred than A29t, although the difference is marginal and equals only 0.05 kcal/mol (see Table 3). This result closely matches the free energy

Table 2

Relative thermodynamic parameters and free energies of solvation (kcal/mol) for DaaPur tautomers in water. For the assignment of symbols see Figure 1. The free energy of solvation ($\Delta\Delta G$) was computed according to Eq. (1). The relative energies shown in parenthesis were computed with BHandHLYP/pVTZ level of theory.

Tautomer	ΔE	ΔH	ΔG	$T\Delta S$	$\Delta\Delta G$
D1	11.30	11.30	11.84	0.55	-16.65
D3	4.91	4.91	5.07	0.16	-10.33
D7	6.09	6.09	6.66	0.57	-3.81
D8	2.59 (3.63)	2.59 (3.63)	2.82 (3.86)	0.14 (-0.23)	-1.22 (-1.30)

Positive values are related to less preferred (relative thermodynamic parameters) or worse solvated (free energy of solvation) tautomers than reference D9 form. Thermodynamics was calculated at 298.15 K.

Table 3

The relative thermodynamic parameters (kcal/mol) for z8isoGua tautomers in the gas phase. For the assignment of symbols see Figure 2. The relative energies shown in parenthesis were computed with BHandHLYP/pVTZ level of theory.

Tautomer	ΔE	ΔH	ΔG	$T\Delta S$
A17	21.58	21.58	21.49	0.09
A19	5.70 (6.06)	5.70 (6.06)	5.32 (6.01)	0.38 (0.06)
A27c	10.87	10.87	11.39	-0.52
A27t	9.55	9.55	10.09	-0.54
A28c	4.64 (4.24)	4.65 (4.24)	5.13 (4.59)	-0.48 (-0.36)
A28t	3.60 (3.12)	3.60 (3.12)	4.16 (3.53)	0.00 (-0.42)
A29c	0.57 (-1.09)	0.57 (-1.09)	1.02 (-0.77)	-0.46 (-0.32)
A29t	0.58 (-1.13)	0.58 (-1.13)	1.16 (-0.72)	-0.58 (-0.41)
A37	7.82	7.82	7.93	-0.11
A39	3.63 (3.42)	3.63 (3.42)	3.67 (3.46)	0.04 (-0.04)

All values are relative to the A38 tautomer. Positive values mean that the A38 tautomer is favored. Thermochemistry was calculated at 298.15 K.

difference obtained for analogous tautomers of isoGua [17], for which amino-enol *cis* form is favored over *trans* form by 0.2 kcal/mol. An interesting observation is that amino-enol *cis* form is preferred for N(9)-H triazole ring protonation. For N(7)-H and N(8)-H protonations (A27(c,t) and A28(c,t) tautomers) the amino-enol *trans* form is preferred by roughly 1 kcal/mol. This effect is probably caused by electrostatic repulsion between the proton located on triazole ring and the proton of the enol group. The proximity of the N(9)-H proton and the enol group in A29 tautomers causes slight shift of the balance of *cis-trans* enol tautomers towards the former one.

The position of a proton in triazole ring naturally divides z8isoGua tautomeric system into three subgroups. Biologically the most important subgroup is the one with N(9)-H protonation, which can serve as a model of a base bound to the sugar ring in a nucleic acid chain. This important subgroup of tautomers is dominated by amino-enol forms as it was previously determined also for isoGua, from both quantum chemical computations [17] and experimental studies [40,41].

The populations of other tautomers in the gas phase is marginal as their ΔG is more than 4 kcal/mol with respect to A29c/t. Tautomers protonated at 7th position (A17, A27, A37) are especially disfavored.

Similarly to DaaPur molecule the tautomeric interconversions of z8isoGua in gas phase are isoentropic ($T\Delta S < 0.5$ kcal/mol) and weakly coupled to the temperature (thermodynamics corrections < 1.0 kcal/mol, see Table S3).

Solvent effect. The A38 tautomer becomes the most stable one in water solution, although the water environment is more favorable for most tautomers except A28(c,t) and A29(c,t) (see last column of Table 4). Especially solvation of A29 tautomers, which have lower energies than A38 in the gas phase, is unfavorable as it increases their free energies by around 5 kcal/mol compared to the solvation of the A38 tautomer. This effect combined with favorable solvation of A39 ($\Delta\Delta G = -1.30$ kcal/mol) and A19 ($\Delta\Delta G = -3.29$ kcal/mol)

makes these two tautomers second and third most populated in water solution, respectively (see Table S4). The population of A39 seems to be higher than A19 as the free energy difference between them is 0.6 kcal/mol. This is an important result, because, N(9)-H tautomers can be treated as models of the part of the nucleotide in a nucleic acid chain. Higher population of A39 tautomer means that A19 tautomer, which can form Watson-Crick interface with iso-Cyt nucleotide, must "compete" with N(3)-H tautomer. Considering that A29c tautomer can form three Watson-Crick hydrogen-bonds with Thymine leads to the conclusion that z8isoGua-isoCyt base-pair can be possibly very mutation-prone as it was observed for isoGua-isoCyt base-pair [12,42]. The energetic proximity of amino-oxo form tautomers protonated at positions N(1) and N(3) and methyl-blocked at position N(9) was also observed by Orozco et al. for isoGua molecule [17]. We checked that methylation of A39 and A19 tautomers of z8isoGua at the position N(9) does not significantly change their relative balance as the free energy difference between them increased only slightly to 0.4 kcal/mol for the lower-level theory computations.

The free energy differences between the lowest (A29c in gas phase and A38 in water) and the second lowest (A38 in gas phase and A39 in water) energy tautomeric forms increased by around 3 kcal/mol, compared to the gas phase calculations, but the difference between lowest and highest energy tautomer (A17) decreased from 22.26 kcal/mol to 10.96 kcal/mol. Although the free energy of solvation of A17 tautomer is very favorable ($\Delta\Delta G = -10.53$ kcal/mol) its population remains very low. In the case of the N(9)-H tautomers the order of keto and enol forms is reversed compared to gas phase, as it was observed for isoGua molecule [17,40,41]. Free energy differences between *cis* and *trans* forms of enol tautomers are less than 0.31 kcal/mol, so the difference significantly decreased compared to the gas phase. Moreover, the order of A28 tautomers was reversed. This observation leads to the conclusion that *cis-trans* balance of amino-enol tautomers is correlated with the proximity of the proton located on triazole ring.

Table 4

The relative thermodynamic parameters and free energies of solvation (kcal/mol) for z8isoGua tautomers in water. For the assignment of symbols see Figure 2. The free energy of solvation ($\Delta\Delta G$) was computed according to Eq. (1). The relative energies shown in parenthesis were computed with BHandHLYP/pVTZ level of theory.

Tautomer	ΔE	ΔH	ΔG	$T\Delta S$	$\Delta\Delta G$
A17	10.25	10.25	10.96	-0.70	-10.53
A19	2.10 (2.17)	2.10 (2.17)	2.97 (2.72)	-0.87 (-0.55)	-2.35 (-3.29)
A27c	9.58	9.58	10.36	-0.78	-1.03
A27t	9.35	9.35	10.05	-0.70	-0.04
A28c	7.28	7.28	8.42	-1.14	3.29
A28t	7.45	7.44	8.65	-1.20	4.49
A29c	4.96 (3.66)	4.96 (3.66)	6.01 (4.39)	-1.05 (-0.73)	4.99 (5.16)
A29t	5.20 (3.83)	5.20 (3.83)	6.30 (4.53)	-1.10 (-0.71)	5.14 (5.25)
A37	5.11 (5.51)	5.11 (5.51)	5.89 (5.95)	-0.78 (-0.44)	-2.04
A39	2.36 (1.87)	2.36 (1.87)	3.01 (2.15)	-0.65 (-0.28)	-0.66 (-1.30)

The reference is the A38 tautomer. Positive values mean that the A38 is more preferred (relative thermodynamic parameters) or better solvated (free energy of solvation). Thermochemistry was calculated at 298.15 K.

Table 5

The vertical excitation energies (eV) for the first three singlet excited states (S1, S2, S3), of all examined tautomers of z8isoGua and DaaPur in water.

Tautomer	Vertical states			Experiment		
	S1	S2	S3	S1	S2	S3
A17	4.295 (4.413) [0.187]	5.334 (5.482) [0.003]	6.289 (6.464) [0.001]			
A27c	5.062 (4.413) [0.234]	5.564 (4.851) [0.004]	5.976 (5.210) [0.009]			
A27t	5.078 (4.413) [0.234]	5.674 (4.930) [0.003]	5.943 (5.163) [0.009]			
A37	5.060 (4.413) [0.223]	5.871 (5.120) [0.008]	6.037 (5.265) [0.253]			
A28c	5.094 (4.413) [0.263]	5.702 (4.941) [0.000]	6.044 (5.237) [0.082]			
A28t	5.104 (4.413) [0.261]	5.684 (5.068) [0.000]	6.037 (5.218) [0.091]			
A38	5.303 (4.413) [0.271]	5.935 (4.939) [0.003]	5.962 (4.961) [0.265]	4.413 ^a		4.960 ^a
A19	4.940 (4.413) [0.223]	5.635 (5.034) [0.008]	6.065 (5.418) [0.222]			
A29c	5.526 (4.413) [0.338]	5.774 (4.611) [0.007]	6.087 (4.861) [0.008]			
A29t	5.530 (4.413) [0.335]	5.788 (4.619) [0.008]	6.087 (4.857) [0.007]			
A39	5.564 (4.413) [0.313]	6.019 (4.774) [0.013]	6.225 (4.937) [0.093]			
D1	4.935 (4.429) [0.185]	5.296 (4.754) [0.006]	5.676 (5.095) [0.031]			
D3	5.334 (4.429) [0.324]	5.649 (4.692) [0.009]	5.854 (4.862) [0.073]			
D7	4.820 (4.429) [0.188]	5.360 (5.175) [0.001]	5.888 (5.412) [0.011]			
D8	4.872 (4.429) [0.230]	5.611 (5.101) [0.000]	5.899 (5.363) [0.111]			
D9	5.304 (4.429) [0.318]	5.801 (4.842) [0.006]	5.978 (4.990) [0.036]	4.429 ^b		4.960 ^b

The optimized ground state energies are the reference values. The scaled energy values and oscillator strengths are given in parenthesis and brackets, respectively.

^a Experimental absorption peaks of z8isoGua, see Ref. [13] (S1) and Ref. [15] (S3).^b Experimental absorption peaks of DaaPur, see Ref. [14].

The smaller distance between enol group and triazole ring proton favors *cis* tautomer, especially in the gas phase. The strength of this effect is reduced by the presence of the polar solvent (water).

Analysis of triazole ring subgroups (as defined for gas phase calculations) revealed that the lowest free energy in each group are A37, A38 and A39 (very closely followed by A19), respectively. Therefore, the pyrimidine ring of free z8isoGua molecule is most probably in N(3)-H tautomeric form, but z8isoGua molecule bound to sugar ring should exist in an equilibrium of N(1)-H and N(3)-H tautomers, most probably shifted towards the latter tautomer. The comparison of computed excitation energies with absorption spectra, also strongly suggests N(3)-H protonation of the pyrimidine ring (see Section 3.3).

Similarly as in the gas phase the tautomeric conversions in water are isotropic (all $T\Delta S$ values are lower than 1.20 kcal/mol), but maximum entropy correction slightly increased. In the case of relative thermodynamic corrections (see Table S4), the difference in free energy also increased and reached 2.2 kcal/mol, which means that correlation between temperature and tautomeric conversions in water is stronger than in the gas phase.

The picture of tautomeric equilibrium presented in this publication is a complement to the information about z8isoGua obtained from the fluorescence emission studies [13]. The only conclusion obtained from fluorescence experiment, which is solely related to the ground state, is a minority of N(8)-H tautomer with respect to other triazole ring tautomers N(7)-H and N(9)-H. This conclusion cannot be confirmed by quantum mechanical computations, as the energetic order of tautomers in water (A38 < A39 < A19 < A29 < A37 < A28 < A27 < A17) favors N(8)-H forms over N(9)-H forms by 2 kcal/mol, although it should be stressed that conclusions drawn from experimental data were semi-quantitative. Further computations and analysis is necessary to figure out the reason of this discrepancy.

3.3. Excited states

The first three singlet excitation energies of each tautomeric form, in water and corresponding oscillation strengths, for both z8isoGua and DaaPur, are presented in Table 5. Almost all calculated excitation energies are higher than those determined experimentally, therefore it is only possible to compare the qualitative trends between computed and experimentally observed transition energies. Excitation energies of each tautomer were scaled as described

in Section 2. The scaling factors range from 0.973 to 1.261. For both DaaPur and z8isoGua molecules the computed and scaled excitation energies in water solution were compared to available experimental data. The excitation energies of investigated compounds in the gas phase are shown in Table S5 in *Supplementary Information*. These energies were not scaled, because absorption spectra are not available for the gas phase.

For DaaPur molecule immersed in water the energy corresponding to the third absorption peak clearly matches scaled excitation energy (S3) of the most populated D9 tautomeric form [14]. For z8isoGua molecule the value of scaled energy of third singlet excitation (S3) which best matches available absorption spectra [13,15], corresponds to A38 tautomer, although the excitation energy of A39 tautomer is also close to the experimental value. This result is consistent with results of ground-state computations as tautomer responsible for absorption spectra is the most abundant one in aqueous solution. Therefore it can be concluded that agreement of positions of maxima of absorption spectra and scaled excitation energies is satisfactory. Moreover, comparison of computed excitation energies and experimental data also strongly suggests N(3)-H protonation of pyrimidine ring.

4. Conclusions

In this publication tautomeric equilibria of amino forms of DaaPur and z8isoGua molecules, both in the gas phase and in water, are investigated by quantum chemistry methods. The tautomeric equilibrium was revealed by the ground-state DFT computations and excitation energies were determined by the application of TDDFT method. The latter results were compared with available experimental data.

The energetic order of DaaPur tautomers in the gas phase (D9 < D8 < D7 < D3 < D1) is only slightly changed by the presence of solvent (D9 < D8 < D3 < D7 < D1). The probability of protonation of DaaPur molecule seems to be correlated with distances between the tautomeric proton and two amino groups. The most stable tautomer, both in water and in the gas phase, is protonated at the position N(9). This result, in agreement with experimental data [14], supports the reliability of results obtained with the same methodology for z8isoGua molecule.

The order of neutral amino tautomers of z8isoGua in the gas phase is A29 < A38 < A39 < A28 < A19 < A37 < A27 < A17. Reordering of tautomers caused by the presence of water (A38 < A39 <

A19 < A29 < A37 < A28 < A27 < A17) is mainly an effect of relatively poor solvation of amino-enol tautomers, for which energies significantly increased. Amino-enol A29 is the most stable tautomer in the gas phase, whilst amino-oxo A38 is the most stable tautomer in water. The free energy difference between biologically most important tautomers A19 and A39 is small and the balance seems to be slightly shifted towards A39 tautomer. This result shows similar behavior of the Watson–Crick interface to the one observed for isoGua [17]. Combined DFT–Poisson–Boltzmann study showed different tautomeric picture of Guanine [43], for which the A19 form is more stable than A39, although the difference is only 1.0 kcal/mol. Nevertheless this difference might be crucial for explanation of the higher specificity of Gua–Cyt compared to isoGua–isoCyt base-pair. The same order of N(9)–H tautomers was also determined by the combination of MP2 method and COSMO solvation model of Guanine [44], although the difference between N(1)–H and N(3)–H tautomers was as big as 9 kcal/mol. As the energetic order of A19–A39 tautomers is reversed for both isoGua and z8isoGua (compared to Gua) the effect of isoCyt → T mutations observed for the expanded genetic code [42] can also be expected for the expanded genetic code in which isoGua is replaced by the fluorescent probe – z8isoGua.

The scaled energies of excitation show good agreement with available experimental data. For z8isoGua the tautomer for which the absorption spectrum best matches scaled excitation energies is A38 – the most abundant tautomer in water. This result, combined with population analysis in the ground state, strongly suggests N(3)–H protonation of pyrimidine ring – an improtant conclusion, which cannot be directly drawn from available experimental data.

Acknowledgments

Authors would like to thank Prof. Jacek Wierchowski, Dr. Hab Bogdan Smyk and Dr. Grzegorz Mędra for valuable discussions and anonymous reviewers for their comments and suggestions, which led to improvement of the manuscript. This work was supported by the Department Internal Grant nos. 528-0706-0802 and 528-0706-0883. M.M. was supported by the European Research Council (ERC grant RNA+P=123D to Janusz M. Bujnicki). All calculations were performed in Regional Computer Center in Olsztyn, CI TASK in Gdańsk, WCSS in Wrocław (Grant no. 349) and ICM KDM in Warsaw (Grant no. G59-14), Poland.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cplett.2015.03.029.

References

[1] J.D. Watson, F.H.C. Crick, *Nature* 171 (1953) 737.

- [2] W. Seanger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, 1984.
- [3] J.S. Kwiatkowski, J. Leszczynski, *J. Mol. Struct.: THEOCHEM* 208 (1–2) (1990) 35.
- [4] M.T. Chenon, R.J. Pugmire, D.M. Grant, R.P. Panzica, L.B. Townsend, *J. Am. Chem. Soc.* 97 (16) (1975) 4636.
- [5] A.O. Alyoubi, R.H. Hilal, *Biophys. Chem.* 55 (3) (1995) 231.
- [6] J. Lin, C. Yu, S. Peng, I. Akiyama, K. Li, L.K. Lee, P.R. LeBreton, *J. Am. Chem. Soc.* 102 (14) (1980) 4627.
- [7] G. Fogarasi, *Chem. Phys.* 349 (1–3) (2008) 204.
- [8] D. Kosenkov, Y. Kholod, L. Gorb, O. Shishkin, D.M. Hovorun, M. Mons, J. Leszczynski, *J. Phys. Chem.* 113 (17) (2009) 6140.
- [9] M. Sabio, S. Topiol, W.C. Lumma, *J. Phys. Chem.* 94 (4) (1990) 1366.
- [10] H.-S. Kim, D.-S. Ahn, S.-Y. Chung, S.K. Kim, S. Lee, *J. Phys. Chem.* 111 (32) (2007) 8012.
- [11] C. Plutzer, K. Kleinermanns, *Phys. Chem. Chem. Phys.* 4 (2002) 4877.
- [12] D. Roberts, R. Bandaru, C. Switzer, *J. Am. Chem. Soc.* 119 (1997) 4640.
- [13] J. Wierchowski, G. Medza, J. Sepiol, M. Szabelski, D. Shugar, *J. Photochem. Photobiol. A* 237 (2012) 64.
- [14] J. Wierchowski, G. Medza, M. Szabelski, A. Stachelska-Wierchowska, *J. Photochem. Photobiol. A* 265 (2013) 49.
- [15] G. Medza, *Spectroscopic Properties and Tautomerization of Chosen 8-Azapurines*, University of Warmia and Mazury, 2014 (Ph.D. thesis).
- [16] E.D. Raczynska, M. Makowski, K. Zientara-Rytter, K. Kolczyńska, T.M. Sępniewski, M. Hallman, *J. Phys. Chem. A* 117 (2013) 1548.
- [17] J.R. Blas, F.J. Luque, M. Orozco, *J. Am. Chem. Soc.* 126 (2004) 154.
- [18] M. Piacenza, S. Grimme, *J. Comput. Chem.* 25 (2004) 83.
- [19] E.D. Raczynska, K. Kolczyńska, B. Ośmiałowski, R. Gawinecki, *Chem. Rev.* 105 (2005) 3561.
- [20] E.D. Raczynska, K. Kolczyńska, *J. Mol. Model.* 18 (2012) 3523.
- [21] E.D. Raczynska, *Comput. Theor. Chem.* 993 (2012) 73.
- [22] E.D. Raczynska, B. Kamińska, *J. Phys. Org. Chem.* 23 (2010) 823.
- [23] E.D. Raczynska, T.M. Sępniewski, K. Kolczyńska, *J. Mol. Model.* 17 (2011) 3229.
- [24] G.R. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
- [25] A.D. Becke, *J. Chem. Phys.* 98 (1993) 5648.
- [26] C. Lee, W. Yang, R.G. Parr, *Phys. Rev. B* 37 (1988) 785.
- [27] W.J. Hehre, L. Radom, P.v R. Schleyer, J.A. Pople, *J. Wiley, Ab Initio Molecular Orbital Theory*, Wiley, 1986.
- [28] S. Miertus, J. Tomasi, *Chem. Phys.* 65 (1982) 239.
- [29] S. Miertus, E. Srocco, J. Tomasi, *Chem. Phys.* 55 (1981) 117.
- [30] C.E. Dykstra, G. Frenking, K.S. Kim, G.E. Scuseria, *Theory and Applications of Computational Chemistry: The First Forty Years*, Elsevier, Amsterdam, 2005.
- [31] M.W. Schmidt, K.K. Baldrige, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, K.A. Nyugen, S. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, *J. Comput. Chem.* 14 (1993) 1347.
- [32] A.D. Becke, *J. Chem. Phys.* 98 (1993) 1372.
- [33] T.H. Dunning, *J. Chem. Phys.* 90 (1989).
- [34] O. Huertas, J.R. Blas, I. Soteras, M. Orozco, F.J. Luque, *J. Phys. Chem. A* 110 (2006) 5010.
- [35] E.K.U. Gross, J.F. Dobson, M. Petersilka, *Top. Curr. Chem.* 181 (1996) 81.
- [36] M.K. Shukla, S.K. Mishra, A. Kumar, P.C. Mishra, *J. Comput. Chem.* 21 (2000) 826.
- [37] M.K. Shukla, P.C. Mishra, *Chem. Phys.* 240 (1999) 319.
- [38] M.K. Shukla, P.C. Mishra, *Chem. Phys.* 230 (1998) 187.
- [39] G.S. Li, M.F. Ruiz-Lopez, M.S. Zhang, B. Maigret, *J. Mol. Struct.: THEOCHEM* 422 (1998) 197.
- [40] J. Sepiol, Z. Kazimierzczuk, D. Shugar, *Z. Naturforsch. C* 31 (1976) 361.
- [41] F. Seela, C. Wei, Z. Kazimierzczuk, *Helv. Chim. Acta* 78 (1995) 1843.
- [42] I. Hirao, *Curr. Opin. Chem. Biol.* 10 (2006) 622.
- [43] Y.H. Jang, K.T. Noyes, L.C. Sowers, S. Hwang, D.S. Chung, *J. Phys. Chem. B* 107 (2003) 344.
- [44] M. Hanus, F. Ryjaček, M. Kabela, T. Kubar, T.V. Bogdan, S.A. Trygubenko, P. Hobza, *J. Am. Chem. Soc.* 125 (2003) 7678.



Contents lists available at ScienceDirect

Computational and Theoretical Chemistry

journal homepage: www.elsevier.com/locate/comptc

Theoretical investigations of tautomeric equilibrium of 9-methyl-8-aza-iso-Guanine and its electrostatic properties

Maciej Pyrka^a, Maciej Maciejczyk^{a,b,*}^a Department of Physics and Biophysics, University of Warmia and Mazury in Olsztyn, ul. Oczapowskiego 4, 10-719 Olsztyn, Poland^b Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, ul. Ks. Trojdena 4, 02-109 Warsaw, Poland

ARTICLE INFO

Article history:

Received 18 April 2016

Received in revised form 25 June 2016

Accepted 1 July 2016

Available online 2 July 2016

Keywords:

Quantum chemistry

Composite methods

Hybrid methods

Tautomerization

Expanded genetic code

Hirshfeld partial charge

Zwitterion

ABSTRACT

Tautomeric equilibrium of 9-methyl-8-aza-iso-Guanine is examined by quantum chemistry hybrid (BHandHLYP/cc-pVTZ) and composite (G3, G4) methods. The most predominant tautomers are the amino-oxo forms protonated at positions N(1) and N(3). This order of domination is consistent with tautomeric populations computed for 8-aza-iso-Guanine, but methylation of the investigated molecule causes significant increase of populations of amino-enol forms protonated at position O(2). The analysis of electrostatic properties of 9-methyl-8-aza-iso-Guanine shows that dominant tautomers have uniform charge distribution, but the vast majority of poorly-populated forms protonated at triazole ring are zwitterions.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The Watson-Crick interface is crucial for canonical base-pairing in nucleic acids. It was shown that tautomeric conversions of nucleobases may change protonation of Watson-Crick interface and consequently cause modification of recognition mechanism and provoke genomic mutations [1]. Many theoretical and experimental studies regarding tautomeric equilibria of nucleobases and alike compounds confirm importance of this phenomenon [2–10].

Development of the expanded genetic code [11] raised the question about tautomeric equilibrium of its new letters – isoGua and isoCyt. The problem was particularly important, because the experimental data showed possibility of isoCyt → Thy mutations [12]. The tautomeric equilibrium of methylated isoGua was revealed by quantum chemistry methods [13]. The methyl group was placed at the position 9 of isoGua in order to mimic its presence in a polynucleotide chain. On the other hand it was shown that derivative of isoGua, in which carbon at position 8 was replaced by nitrogen, is a promising fluorescence probe [14,15]. The tautomeric equilibrium of 8-aza-iso-Gua (z8isoGua) was recently investigated in our laboratory [16].

In this paper tautomeric equilibrium of 9-methyl derivative of z8isoGua (9m-z8isoGua) is examined by quantum chemistry methods. The methyl group mimics deoxyribose ring, therefore the investigated system emulates z8isoGua molecule inserted in a polynucleotide chain. 9m-z8isoGua has 12 more possible tautomers compared to methylated isoGua, which is caused by more extensive protonation of triazole ring than in the case of the latter compound and this in turn may affect the electronic properties of 9m-z8isoGua. Therefore, Hirshfeld charge distribution analysis was also performed. Computations were carried out at the same level of the quantum theory as those applied to z8isoGua [16] and compared to the results obtained with composite Gaussian-3 [17] and Gaussian-4 [18] methods.

2. Methods

Ab initio methods were applied to examine tautomeric nature of nucleobases and its analogs [13,19,20]. The geometry optimization, vibrational analysis and thermochemistry calculations were performed by using GAMESS package with application of DFT-BHandHLYP method [21] and cc-pVTZ basis set [22]. It was recently suggested [16] that BHandHLYP/cc-pVTZ level of theory should be sufficient for a proper study of z8isoGua tautomeric equilibrium. Nevertheless validity of this methodology was checked by comparison of all results to those obtained with

* Corresponding author at: Department of Physics and Biophysics, University of Warmia and Mazury in Olsztyn, ul. Oczapowskiego 4, 10-719 Olsztyn, Poland.
E-mail address: maciej.maciejczyk@uwm.edu.pl (M. Maciejczyk).

high-level composite Gaussian-3 (G3) method [17], which is a good compromise between a speed of computations and an accuracy of the results. Free energies of low energy tautomers were also computed with Gaussian-4 (G4) method [18]. In order to reveal electronic properties of investigated system, calculations of Hirshfeld partial charges [23] were performed with Gaussian 03 package. Hirshfeld partial charges were previously computed also for methyl derivatives of guanine [24], which binds to the transcription initiation protein [25]. All possible tautomeric forms of 9m-z8isoGua were optimized in the gas phase and vibrational analysis was performed. All tautomers computed with BHandHLYP/cc-pVTZ method reached stationary point and therefore thermochemistry calculations were performed. Procedure was repeated in water solution emulated by PCM method [26,27]. The final structures obtained from optimization in vacuum were used as the initial conformations for optimization in water.

The total electric dipole moment and Hirshfeld partial charges were used to examine electronic charge distribution of all tautomers. The special emphasis was put on detection of zwitterionic forms, which were defined in the following manner. For each tautomer partial charges of triazole and pyrimidine rings were summed up (Q_T and Q_P respectively). C(4) and C(5) carbons were excluded from the summation (for the atom numbering see tautomer AO1 in Fig. 1). If the absolute value of both Q_T and Q_P reaches at least 0.3 a.u. then the tautomer is defined as zwitterion. Note that neutrality of investigated systems practically guarantees opposite charges of both rings for zwitterionic molecules. The same analysis was performed for Mulliken and Löwdin partial charges and no significant qualitative differences were observed, in particular exactly the same tautomers were predicted as zwitterions.

Gibbs free energy ($G = H - TS$), energy (E – including electronic energy and zero point energy), enthalpy (H) and entropy (S), determined from thermochemistry computations, were used for computations of population distribution. Computed parameters also show dependence of tautomeric system on temperature and entropy. The relative Gibbs free energies (ΔG) of each tautomer with respect to AO1 tautomer were computed. Later they were used to determine the percentage contents of individual tautomers ($x_i = \frac{\Delta G_i}{\sum_{j=1}^n \Delta G_j} \cdot 100\%$). ΔG includes following terms: electronic energy, zero point energy and thermal corrections to energy and entropy. Other relative thermodynamic parameters were also determined ($\Delta E, \Delta H, T\Delta S$). The solvation effect of tautomers was examined by calculation of solvation free energy $\Delta\Delta G$ in a following way:

$$\Delta\Delta G = \Delta G_{\text{water}} - \Delta G_{\text{gas}} \quad (1)$$

where ΔG_{water} and ΔG_{gas} are relative free energies of a given tautomer in water and vacuum, respectively.

Figs. 1–3 depict all 34 possible tautomeric forms of 9m-z8isoGua. Tautomers are divided into families according to their donor-acceptor pattern of the Watson-Crick interface. The biologically most important tautomeric forms, which can form three hydrogen bonds with either IsoCyt (left column) or Thy (right column), are shown in Fig. 1. Other possible tautomeric forms are depicted in Figs. 2 and 3.

3. Results and discussion

3.1. Comparison of quantum chemistry methods

Relative Gibbs free energies for all tautomers in the gas phase and in water were computed with hybrid BHandHLYP/cc-pVTZ and composite Gaussian-3 methods (see Tables 1 and 2). The

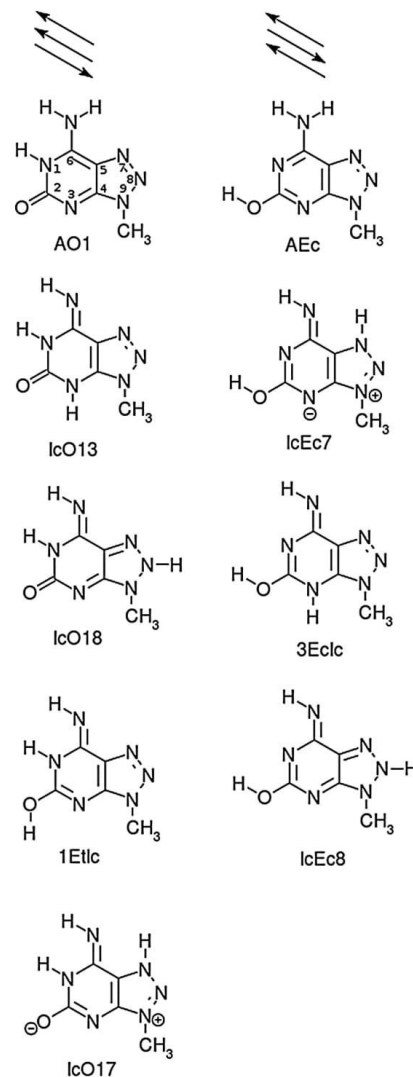


Fig. 1. Biologically important tautomers which can form Watson-Crick hydrogen bonds with either IsoCytosine (left) or Thymine (right). The arrows indicate that molecule is either donor or acceptor of a proton in a hydrogen bond. The atom numbering is shown on AO1 tautomer.

results appear to be highly correlated as can be seen in Fig. 4, although free energy differences obtained with G3 method are smaller. Very good correlation between results obtained with vastly different quantum chemistry methods strongly supports validity of the tautomeric order predicted by BHandHLYP/cc-pVTZ method and the results obtained with this method were used in the further analysis. It should also be noticed that free energy differences of four lowest-energy tautomers (AO1, AO3, AEc(t)) in water obtained with BHandHLYP/cc-pVTZ method almost perfectly

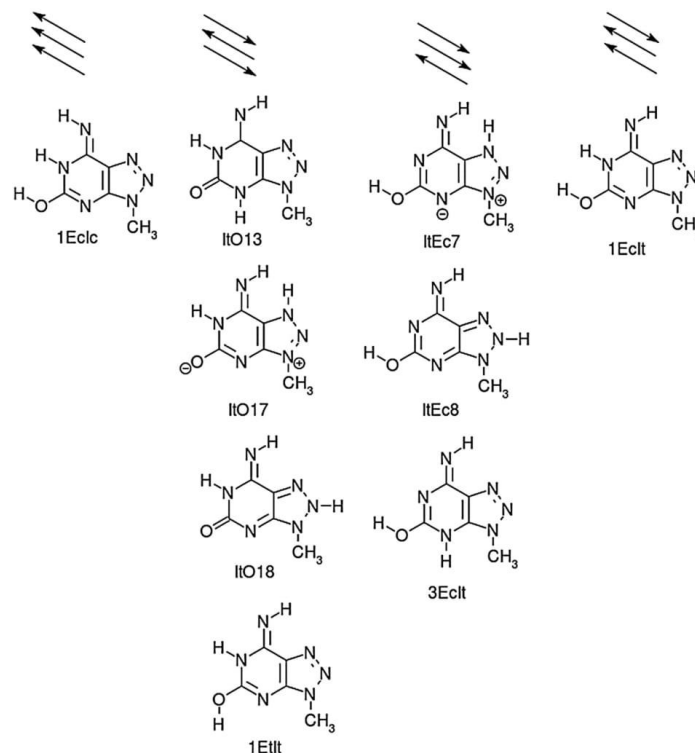


Fig. 2. Tautomers which have a low tendency to form Watson-Crick interface with IsoCytosine and Thymine due to unfavorable protonation pattern. Arrows have the same meaning as in Fig. 1.

match those obtained for isoGua with top-level QM methods – MP4/aug-cc-pVTZ and CCSD(T)/aug-cc-pVTZ (see Table 3 of Ref. [13]). Free energy differences computed with G4 method for low-energy tautomers do not significantly differ from those obtained with G3 (see Tables 1 and 2).

3.2. 9m-z8isoGua tautomerism

Replacement of carbon at position 8 by nitrogen in isoGua severely increases the number of possible tautomers, because in z8isoGua (or 9m-z8isoGua) position 8 can be either protonated or deprotonated. In isoGua C(8) is protonated for all possible tautomers [13]. The relative Gibbs free energies (ΔG) as well as relative energies (ΔE), enthalpies (ΔH) and entropic contributions ($T\Delta S$) obtained from thermochemistry computations in vacuum and in water solution are presented in Tables 1 and 2, respectively.

Gas phase. The data collected in Table 1 show that amino-enol forms, AEc and AEt, are the most favored tautomers. The difference between *cis* and *trans* form is less than 1 kcal/mol in favor of the first one. This fact may be caused by the electrostatic repulsive interaction decaying with the distance between the (O2)-H and the methyl group, although it should be stressed that the difference is small, most probably below the uncertainty of BHandHLYP method. The domination of amino-enol tautomers in vacuum was also shown in theoretical and experimental findings of isoGua [13,28,29] and computational studies of z8isoGua [16]. This coinci-

dence supports our suggestion that application of BHandHLYP/cc-pVTZ level of theory should be an appropriate approach.

Populations of other tautomeric forms are negligible. The free energies of other tautomers (including biologically important forms – AO1, AO3 and ItO13) are at least 4 kcal/mol larger than the free energy of AEc. It should be noticed that in three subgroups (amino-oxo, imino-oxo, imino-enol), practically all forms protonated at position N(7) and N(8) are the most discriminated in a given subgroup. Similarly as in the case of *cis/trans* explanation, the reason is a relatively short distance between proton located at positions 7 or 9 and the methyl group.

The tautomeric interconversions have isoentropic character, similarly as it was the case for z8isoGua and similar molecules [16,19,30–35]. Results presented in Table 1 show that clear majority of $T\Delta S$ values are lower than 1 kcal/mol. However, roughly thirty percent of the values of thermodynamic correction to ΔE , ΔH , ΔG (especially for the last one) is larger than 1 kcal/mol. This means, that tautomeric interconversion process for 9m-z8isoGua is softly temperature-dependent.

Solvent effect. Immersion of 9m-z8isoGua in an aqueous solution substantially reorganized energetic order of tautomers prevailed in the gas phase. Free energies of solvation obtained with both methods are collected in Table 3. As in the case of isoGua and z8isoGua [13,16], the most favored tautomers are amino-oxo forms protonated at position 1 and 3 (AO1 and AO3) with a small advantage of the first one (0.27 kcal/mol). The electric dipole

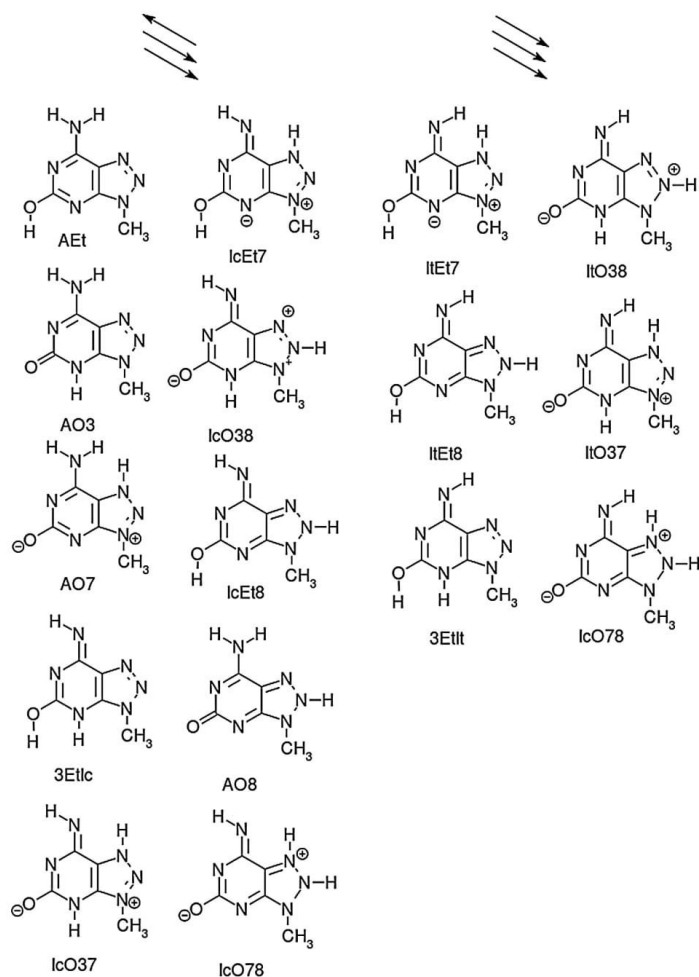


Fig. 3. As Fig. 2, but with different donor-acceptor pattern.

moment p of AO1 is larger than AO3 ($p = 7.79$ D and 5.82 D), therefore the latter one is worse solvated in a polar solution ($\Delta\Delta G = 2.84$ kcal/mol). The order of these tautomers is reversed compared to the gas phase results and the absolute difference is significantly decreased (from $|\Delta G| = 2.57$ kcal/mol in gas phase to 0.27 kcal/mol in water), therefore the populations of these two forms are similar (see Table S2 of SI). It was previously observed that analogical amino-oxo tautomers of isoGua are able to form Watson-Crick interface with isoCyt [11,12]. Therefore, exactly as it was the case for isoGua and z8isoGua energetic vicinity of AO1 and AO3 tautomers causes the effect of their competition for the favor of isoCyt.

Amino-enol forms AEt(c), despite the worst solvation ($\Delta\Delta G = 6.91$ kcal/mol and 8.02 kcal/mol; $p = 0.90$ D and 2.85 D, respectively), are still substantially present in water solution ($\Delta G = 0.90$ kcal/mol in AEt and 1.14 kcal/mol in AEc). Especially,

significant presence of AEc tautomer is very notable as it can form three hydrogen bonds via Watson-Crick interface with Thymine [13]. Most of the imino-oxo tautomers are better solvated than AO1 tautomer, however they are still energetically discriminated. It is hard to find explanation of the relatively good solvation of imino-enol tautomers, although all forms protonated at position N(1) are more difficult to solvate compared to AO1 tautomer. As expected, with the exception of four forms (IcEc7, IcEt7, IcEc8, ItO18), protonation at triazole ring, which causes significant charge separation between the rings, impacts favorably the solvation process. Nevertheless imino-enol tautomers remain very unstable. The influence of solvent causes reduction of the energetic gap between the most favored and most unstable tautomer (about 20 kcal/mol).

Isoentropic character of tautomeric conversions is confirmed by computation of relative entropies (most $T\Delta S < 1$ kcal/mol) and it is similar to the observations obtained from the gas phase computa-

Table 1

The relative thermodynamic parameters (kcal/mol) computed with BHandHLYP/cc-pVTZ method for 9m-z8isoGua tautomers in the gas phase. The reliability of the method was checked by the comparison of computed Gibbs free energies with results obtained with Gaussian-3 and Gaussian-4 methods (last two columns). For the assignment of symbols see Figs. 1–3.

Tautomer	BHandHLYP			Gaussian-3	Gaussian-4
	$\Delta E, \Delta H$	$T\Delta S$	ΔG	ΔG	ΔG
AEc	-7.08	-0.19	-6.88	-7.74	-8.24
AEt	-7.59	-1.58	-6.01	-7.74	-7.51
AO3	-2.69	-0.12	-2.57	-2.84	-2.48
AO7	31.18	-1.44	32.62	27.38	
AO8	29.07	-0.17	29.24	28.00	
1Ec1c	18.34	0.58	17.76	15.49	
1Ect1	11.29	-0.47	11.76	9.49	
1Et1c	10.40	-0.39	10.79	8.70	
1Et1t	3.72	-1.96	5.67	2.98	3.25
3Ec1c	20.52	-0.32	20.84	18.16	
3Ect1	20.43	-0.30	20.73	18.19	
3Et1c	29.86	-0.27	30.13	26.89	
3Et1t	30.54	0.23	30.77	27.14	
1cEc7	24.66	-2.13	26.80	21.81	
1cEt7	26.97	-0.54	27.51	22.44	
1tEc7	33.41	-2.07	35.49	29.50	
1tEt7	36.50	-0.51	37.01	31.72	
1cEc8	34.51	-0.74	35.25	33.65	
1cEt8	36.36	-0.71	37.07	35.18	
1tEc8	37.19	-0.72	37.92	35.83	
1tEt8	39.87	-0.64	40.52	38.11	
1cO13	4.18	-0.39	4.58	3.96	4.03
1tO13	-0.78	-0.43	-0.34	-0.89	-0.61
1cO17	18.86	-1.87	20.73	16.99	
1tO17	23.43	-0.59	24.02	20.40	
1cO18	25.10	-0.68	25.78	25.21	
1tO18	23.08	-0.64	23.72	23.01	
1cO37	38.21	-0.27	38.49	34.31	
1tO37	50.17	-0.06	50.23	45.14	
1cO38	50.48	-0.44	50.92	-	
1tO38	55.48	-0.38	55.86	53.12	
1cO78	85.06	-0.43	85.48	80.93	
1tO78	99.32	0.04	99.28	93.64	

All values are relative to the A01 tautomer. Positive values mean that the A01 tautomer is favored. Thermochemistry was calculated at 298.15 K. G3 computations of 1cO38 tautomer could not converge.

tions. Whereas, inclusion of water increased coupling between the investigated system and the temperature, since amount of relative thermal corrections, which exceed 1 kcal/mol, increased from roughly thirty to nearly fifty percent.

Partial charge distribution. The electric dipole moments (p) and sum charges (Q_p, Q_r) for each ring of the individual tautomer are given in Table 4. The electric dipole moment of the most populated tautomers both in the gas phase (AEc(t)) and in water (AO1, AO3) do not exceed 8 D. Furthermore, partial charge distribution of these most favored molecules is similar, namely Q_p and Q_r values oscillates around zero. In contrary to above observation, tautomers protonated at positions N(7) or N(8) mostly exhibit a high value of electric dipole moment (from 7.03 D in 1cEc7 to 23.74 D in 1tO78). Moreover, rings of only these specific forms are subject of unequal charge distribution (from -1.07 a.u. on the pyrimidine ring of 1tO78 to 0.76 a.u. on the triazole ring of 1cO78). Interestingly, a negative charge is accumulated on the pyrimidine ring and positive charge on triazole ring.

The investigated tautomeric system includes several tautomers which can be defined as zwitterions. These molecules, namely AO7, 1cEc(t)7, 1c(t)O17, 1c(t)O37, 1c(t)O38 and 1c(t)O78, are marked with italic type in Table 4. The zwitterionic character of these molecules can be deduced from simple assignment of covalent bond order and this observation was also confirmed by the partial charge distribution analysis. It should be stressed here that it is not possible to assign bond orders in such a way that these molecules are not

Table 2

The relative thermodynamic parameters (kcal/mol) computed with BHandHLYP/cc-pVTZ method for 9m-z8isoGua tautomers in water. The reliability of the method was checked by the comparison of computed Gibbs free energies with results obtained with Gaussian-3 and Gaussian-4 methods (last two columns).

Tautomer	BHandHLYP			Gaussian-3	Gaussian-4
	$\Delta E, \Delta H$	$T\Delta S$	ΔG	ΔG	ΔG
AEc	0.22	-0.92	1.14	0.11	0.06
AEt	0.84	-0.05	0.90	0.19	0.20
AO3	-0.12	-0.34	0.27	0.04	-0.17
AO7	21.67	-0.06	21.73	20.03	
AO8	24.46	-1.09	25.55	22.51	
1Ec1c	18.62	-0.32	18.94	16.38	
1Ect1	15.15	-1.23	16.38	14.31	
1Et1c	13.79	-0.48	14.27	12.64	
1Et1t	11.75	-0.38	12.14	10.80	
3Ec1c	21.83	-0.41	22.25	19.61	
3Ect1	22.42	-0.33	22.74	19.74	
3Et1c	28.05	-1.05	29.10	23.02	
3Et1t	28.65	-1.11	29.76	23.46	
1cEc7	29.34	-0.89	30.24	26.90	
1cEt7	29.21	-1.50	30.71	27.17	
1tEc7	32.90	-0.47	33.37	29.20	
1tEt7	33.35	-0.85	34.20	29.64	
1cEc8	35.07	-0.30	35.37	33.10	
1cEt8	35.66	-0.13	35.78	33.53	
1tEc8	36.15	-0.24	36.39	34.01	
1tEt8	36.72	-0.16	36.88	34.52	
1cO13	6.17	-0.80	6.97	5.71	5.26
1tO13	4.25	-0.79	5.04	4.00	3.66
1cO17	19.14	-0.20	19.33	18.67	
1tO17	20.38	0.01	20.37	20.13	
1cO18	23.90	-0.56	24.46	24.01	
1tO18	23.36	-0.66	24.02	22.91	
1cO37	31.50	-0.83	32.33	28.55	
1tO37	35.57	-0.75	36.32	31.64	
1cO38	37.30	-0.54	38.44	34.77	
1tO38	39.43	-0.55	39.98	36.05	
1cO78	72.94	-0.23	73.17	69.44	
1tO78	77.23	-0.54	77.77	74.27	

The reference is the A01 tautomer. Positive values mean that the A01 is more preferred (relative ΔG). Thermochemistry was calculated at 298.15 K.

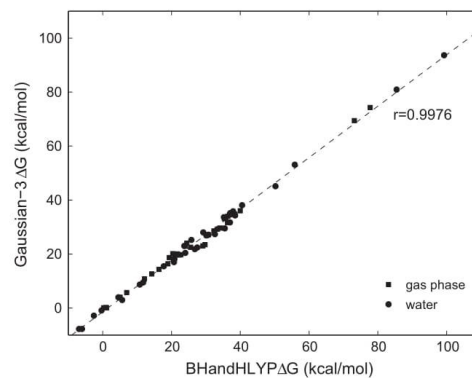


Fig. 4. Correlation between relative Gibbs free energies (ΔG) obtained with BHandHLYP/cc-pVTZ and Gaussian-3 methods. The correlation coefficient is 0.9976.

zwitterions. All zwitterions concluded from bond order analysis have also zwitterionic charge distribution confirmed by Hirshfeld partial charge analysis. Nevertheless there are several tautomers, namely AO8, 1c(t)Ec(t)8 and 1c(t)O18, for which zwitterionic character can only be concluded from the analysis of partial charges.

Table 3

The relative free energies of solvation ($\Delta\Delta G$) computed with BHandHLYP/cc-pVTZ, Gaussian-3 and Gaussian-4 methods (kcal/mol).

Tautomer	BHandHLYP	Gaussian-3	Gaussian-4
AEc	8.02	7.85	8.30
AEt	6.91	7.94	7.70
AO3	2.84	2.88	2.23
AO7	-10.89	-7.35	
AO8	-3.68	-5.49	
1Ec1c	1.18	0.89	
1Ect	4.63	4.81	
1Et1c	3.48	3.94	
1Ett	6.46	7.83	
3Ec1c	1.41	1.45	
3Ect	2.02	1.56	
3Et1c	-1.03	-3.87	
3Ett	-1.01	-3.68	
1cEc7	3.44	5.09	
1cEt7	3.20	4.73	
1tEc7	-2.11	-0.30	
1tEt7	-2.81	-2.07	
1cEc8	0.12	-0.55	
1cEt8	-1.29	-1.65	
1tEc8	-1.53	-1.82	
1tEt8	-3.64	-3.58	
1cO13	2.40	1.75	1.23
1tO13	5.38	4.89	4.27
1cO17	-1.40	1.69	
1tO17	-3.65	-0.28	
1cO18	-1.32	-1.20	
1tO18	0.30	-0.11	
1cO37	-6.16	-5.76	
1tO37	-13.92	-13.50	
1cO38	-12.48	-	
1tO38	-15.89	-17.07	
1cO78	-12.31	-11.49	
1tO78	-12.32	-19.37	

The reference is the AO1 tautomer.

These tautomers can be sketched either in zwitterionic or nonzwitterionic forms as can be seen in Fig. 5 for AO8 tautomer. The only way to pick up the correct form is partial charge analysis of the electronic charge distribution obtained from QM computations.

4. Conclusions

In this work tautomeric equilibrium of 9-methyl derivative of z8isoGua was investigated by hybrid BHandHLYP/cc-pVTZ and composite Gaussian-3 and Gaussian-4 quantum chemistry methods and results were compared to those obtained for methyl-blocked isoGua [13] and nonmethylated z8isoGua molecule [16]. Relative free energies obtained with BHandHLYP and G3 are strongly correlated and results obtained with G4 method do not significantly differ from those obtained with G3. Strong correlation between results obtained with hybrid BHandHLYP/cc-pVTZ and composite G3 methods supports validity of the determined tautomeric order. Based on comparison of our results with top-level QM data from Ref. [13] obtained for isoGua we suggest that free energies obtained with BHandHLYP/cc-pVTZ method are quantitatively more reliable, although both methods predict roughly the same order of tautomers. Therefore we suggest that computationally less demanding G3 method is well suited for prediction of tautomeric order and preliminary selection of low-energy tautomers, but more expensive BHandHLYP/cc-pVTZ method should be used for quantitative analysis.

Similarly as it was the case for m9-isoGua and z8isoGua amino-oxo forms AO1 and AO3 dominate, but surprisingly methylation of z8isoGua significantly increase populations of amino-enol forms AEc(t). This effect is not caused by solvation, but rather by changes in the electronic structure of the molecule. These new highly-

Table 4

The electric dipole moments p (D) and Hirshfeld partial charges Q_p and Q_T located on pyrimidine and triazol ring, respectively. For definition of Q_p and Q_T see Section 2. Zwitterions which can be deduced from simple assignment of covalent bond order are marked italic type and zwitterions which can be deduced only from partial charge distribution are marked bold type.

Tautomer	p	Q_p	Q_T
AEc	2.85	0.002	-0.090
AEt	0.90	-0.003	-0.089
AO1	7.79	0.002	-0.091
AO3	5.82	-0.068	-0.048
AO7	16.94	-0.621	0.490
AO8	17.06	-0.581	0.436
1Ec1c	9.12	0.016	-0.100
1Ect	5.94	0.004	-0.091
1Et1c	9.73	0.012	-0.099
1Ett	7.54	0.003	-0.091
3Ec1c	12.41	-0.038	-0.069
3Ect	11.72	-0.053	-0.060
3Et1c	15.10	-0.040	-0.068
3Ett	14.94	-0.054	-0.058
<i>1cEc7</i>	7.03	-0.633	0.489
<i>1cEt7</i>	10.05	-0.636	0.490
<i>1tEc7</i>	10.42	-0.642	0.497
<i>1tEt7</i>	13.16	-0.643	0.498
<i>1cEc8</i>	10.84	-0.603	0.449
<i>1cEt8</i>	14.29	-0.607	0.452
<i>1tEc8</i>	13.63	-0.612	0.457
<i>1tEt8</i>	16.84	-0.616	0.459
1cO13	7.17	-0.053	-0.057
1tO13	7.45	-0.066	-0.047
<i>1cO17</i>	9.35	-0.626	0.485
<i>1tO17</i>	12.89	-0.637	0.495
<i>1cO18</i>	10.69	-0.594	0.440
<i>1tO18</i>	14.21	-0.604	0.449
<i>1cO37</i>	15.44	-0.729	0.544
<i>1tO37</i>	18.87	-0.740	0.552
<i>1cO38</i>	19.63	-0.720	0.523
<i>1tO38</i>	22.63	-0.731	0.532
<i>1cO78</i>	23.74	-1.049	0.750
<i>1tO78</i>	20.24	-1.066	0.763

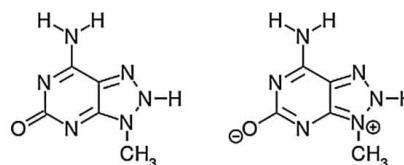


Fig. 5. Nonzwitterionic (left) and zwitterionic (right) forms of AO8 tautomer. Quantum chemistry methods show that the latter form is the correct one.

populated tautomers, as well as dominant AO1 and AO3 forms were also observed for methylated isoGua [13]. As the AEc tautomer has hydrogen-bond donor-acceptor pattern, which may result in a formation of three hydrogen bonds with Thymine, its significant presence may lead isoCyt \rightarrow Thy genomic mutation, similarly as it was the case for isoGua [12]. Imino forms of methylated z8isoGua are discriminated as it was also observed for m9-isoGua and z8isoGua.

The partial charges located on the pyrimidine (Q_p) and triazole (Q_T) ring in combination with results of dipole moments (p) clearly show, that tautomers with protonated triazole ring are zwitterion species. Moreover for several tautomers it is not possible to recognize zwitterionic forms from simple bond-order considerations. Their zwitterionic character can only be confirmed by partial charge analysis of QM-determined charge distribution.

Acknowledgments

This work was supported by the department internal grants: number 528-0706-0802 and 528-0706-0883. All calculations were performed in Regional Computer Center in Olsztyn, Academic Computer Center in Gdańsk, Wrocław Centre for Networking and Supercomputing (<http://wcss.pl>, Grant No. 349) and ICM KDM in Warsaw (Grant No. G59-14), Poland.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.comptc.2016.07.001>.

References

- [1] W. Seanger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, 1984.
- [2] J.S. Kwiatkowski, J. Leszczyński, An ab initio quantum-mechanical study of tautomerism of purine, adenine and guanine, *J. Mol. Struct.: THEOCHEM* 208 (1–2) (1990) 35–44.
- [3] M.T. Chenon, R.J. Pugmire, D.M. Grant, R.P. Panzica, L.B. Townsend, Carbon-13 magnetic resonance. XXVI. Quantitative determination of the tautomeric populations of certain purines, *J. Am. Chem. Soc.* 97 (16) (1975) 4636–4642.
- [4] A.O. Alyoubi, R.H. Hilal, A theoretical and experimental investigation of the electronic spectra and tautomerization of nucleobases, *Biophys. Chem.* 55 (3) (1995) 231–237.
- [5] J. Lin, C. Yu, S. Peng, I. Akiyama, K. Li, L.K. Lee, P.R. LeBreton, Ultraviolet photoelectron studies of the ground-state electronic structure and gas-phase tautomerism of purine and adenine, *J. Am. Chem. Soc.* 102 (14) (1980) 4627–4631.
- [6] G. Fogarasi, Water-mediated tautomerization of cytosine to the rare imino form: an ab initio dynamics study, *Chem. Phys.* 349 (1–3) (2008) 204–209.
- [7] D. Kosenkov, Y. Kholod, L. Gorb, O. Shishkin, D.M. Hovorun, M. Mons, J. Leszczyński, Ab initio kinetic simulation of gas-phase experiments: tautomerization of cytosine and guanine, *J. Phys. Chem.* 113 (17) (2009) 6140–6150.
- [8] M. Sabio, S. Topiol, W.C. Lumma, An investigation of tautomerism in adenine and guanine, *J. Phys. Chem.* 94 (4) (1990) 1366–1372.
- [9] H.-S. Kim, D.-S. Ahn, S.-Y. Chung, S.K. Kim, S. Lee, Tautomerization of adenine facilitated by water: computational study of microsolvation, *J. Phys. Chem.* 111 (32) (2007) 8007–8012.
- [10] C. Plutzer, K. Kleiner, Tautomers and electronic states of jet-cooled adenine investigated by double resonance spectroscopy, *Phys. Chem. Chem. Phys.* 4 (2002) 4877–4882.
- [11] C. Roberts, R. Bandaru, C. Switzer, Theoretical and experimental study of isoguanine and isocytosine: base pairing in an expanded genetic system, *J. Am. Chem. Soc.* 119 (1997) 4640–4649.
- [12] I. Hirao, Unnatural base pair systems for DNA/RNA-based biotechnology, *Curr. Opin. Chem. Biol.* 10 (2006) 622–627.
- [13] J.R. Blas, F.J. Luque, M. Orozco, Unique tautomeric properties of isoguanine, *J. Am. Chem. Soc.* 126 (2004) 154–164.
- [14] J. Wierzchowski, G. Mędza, J. Sepiol, M. Szabelski, D. Shugar, Fluorescent emission properties of 8-azaisoguanine and its N-methyl derivatives: Ground- and excited-state tautomerism, *J. Photochem. Photobiol. A* 237 (2012) 64–70.
- [15] A. Stachelska-Wierzchowska, J. Wierzchowski, A. Bzowska, B. Wielgus-Kutrowska, Site-selective ribosylation of fluorescent nucleobase analogs using purine-nucleoside phosphorylase as a catalyst: effects of point mutations, *Molecules* 21 (2016) 44.
- [16] M. Pyrka, M. Maciejczyk, Theoretical study of tautomeric equilibria of 2,6-diamino-8-azapurine and 8-aza-iso-Guanine, *Chem. Phys. Lett.* 627 (2015) 30–35.
- [17] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rassolov, J.A. Pople, Gaussian-3 (G3) theory for molecules containing first and second-row atoms, *J. Chem. Phys.* 109 (1998) 7764.
- [18] L.A. Curtiss, P.C. Redfern, K. Raghavachari, Gaussian-4 theory, *J. Chem. Phys.* 126 (2007) 084108.
- [19] E.D. Raczynska, M. Makowski, K. Zientara-Rytter, K. Kolczyńska, T.M. Stępniewski, M. Hallman, Quantum-chemical studies on the favored and rare tautomers of neutral and redox adenine, *J. Phys. Chem. A* 117 (2013) 1548–1559.
- [20] M. Piacenza, S. Grimme, Systematic quantum chemical study of DNA-base tautomers, *J. Comput. Chem.* 25 (2004) 83–98.
- [21] A.D. Becke, A new mixing of Hartree-Fock and local density-functional theories, *J. Chem. Phys.* 98 (1993) 1372–1377.
- [22] T.H. Dunning, Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, *J. Chem. Phys.* 90 (1989) 1007–1023.
- [23] E.L. Hirshfeld, Spatial partitioning of charge density, *Israel J. Chem.* 16 (1977) 198–201.
- [24] K. Rusczyńska, K. Kamińska-Trela, J. Wójcik, J. Stepieński, E. Darzyńkiewicz, R. Stolarski, Spatial partitioning of charge density, *Biophys. J.* 85 (2003) 1450–1456.
- [25] K. Rusczyńska-Bartnik, M. Maciejczyk, R. Stolarski, Dynamical insight into *Caenorhabditis elegans* eIF4E recognition specificity for mono- and trimethylated structures of mRNA 5' cap, *J. Mol. Model.* 17 (2011) 727–737.
- [26] S. Miertus, J. Tomasi, Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes, *Chem. Phys.* 65 (1982) 239–241.
- [27] S. Miertus, E. Srocco, J. Tomasi, Electrostatic interaction of a solute with a continuum. A direct utilization of AB initio molecular potentials for the prevision of solvent effects, *Chem. Phys.* 55 (1981) 117–129.
- [28] J. Sepiol, Z. Kazimierzczuk, D. Shugar, Tautomerism of isoguanosine and solvent-induced keto-enol equilibrium, *Z. Naturforsch. C* 31 (1976) 361–370.
- [29] F. Seela, C. Wei, Z. Kazimierzczuk, Substituent reactivity and tautomerism of isoguanosine and related nucleosides, *Helvetica Chim. Acta* 78 (1995) 1843.
- [30] E.D. Raczynska, K. Kolczyńska, B. Ośmiałowski, R. Gawinecki, Tautomeric equilibria in relation to pi-electron delocalization, *Chem. Rev.* 105 (2005) 3561–3612.
- [31] E.D. Raczynska, K. Kolczyńska, T.M. Stępniewski, Consequences of one-electron oxidation and one-electron reduction for 4-aminopyrimidine – DFT studies, *J. Mol. Model.* 18 (2012) 3523–3533.
- [32] E.D. Raczynska, Quantum-chemical studies of the consequences of one-electron oxidation and one-electron reduction for imidazole in the gas phase and water, *Comput. Theor. Chem.* 993 (2012) 73–79.
- [33] E.D. Raczynska, B. Kamińska, Prototropy and pi-electron delocalization for purine and its radical ions – DFT studies, *J. Phys. Organ. Chem.* 23 (2010) 823–835.
- [34] E.D. Raczynska, T.M. Stępniewski, K. Kolczyńska, Consequences of one-electron oxidation and one-electron reduction for aniline, *J. Mol. Model.* 17 (2011) 3229–3239.
- [35] G.S. Li, M.F. Ruiz-Lopez, M.S. Zhang, B. Maigret, Ab initio calculations of tautomer equilibrium and protonation enthalpy of 4- and 5-methylimidazole in the gas phase: basis set and correlation effects, *J. Mol. Struct.: THEOCHEM* 422 (1998) 197–204.

Why Purine Nucleoside Phosphorylase Ribosylates 2,6-Diamino-8-azapurine in Noncanonical Positions? A Molecular Modeling Study

Maciej Pyrka and Maciej Maciejczyk*



Cite This: *J. Chem. Inf. Model.* 2020, 60, 1595–1606



Read Online

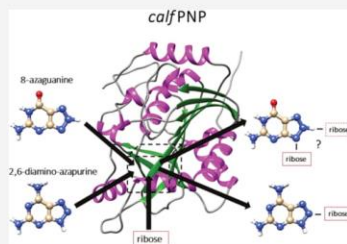
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Protein nucleoside phosphorylase (PNP) is an enzyme that catalyzes a reversible conversion process (ribosylation and phosphorolysis) between nucleobases (purines) and their nucleosides. Experimental studies showed that calf PNP ribosylates purine analogues in specific positions: 2,6-diamino-8-azapurine in position 7 or 8 and 8-azaguanine in position 9 of the triazole ring. The reason for this phenomenon can be a result of different expositions of purine substrates to the channel leading to the binding site. This hypothesis was verified by the application of molecular modeling techniques to two complexes of purine analogues 2,6-diamino-azapurine, calf PNP (pdb-code: 1LVU), and 8-azaguanine, calf PNP (pdb-code: 2A11). The results obtained with a combination of quantum chemistry, docking, and molecular dynamics methods showed qualitative validity of our hypothesis. Binding free energies of protein–ligand systems showed that most probable binding poses expose N8 nitrogen for 2,6-diamino-8-azapurine and N9 nitrogen for 8-azaguanine into the binding channel and ruled out the exposition of N9 for 2,6-diamino-8-azapurine and N7 for 8-azaguanine, partially in agreement with the experimental data. The other important result obtained in this study is a significantly higher population of the protonated form of crucial residue Glu-201 present in the binding pocket, compared to the standard protonation of free glutamic acid in solution. This result combined with populations of tautomeric forms of both investigated systems strongly suggests that 2,6-diamino-8-azapurine and 8-azaguanine are recognized by proteins with deprotonated and protonated Glu-201 residues, respectively. A comparison of computed binding poses of the investigated ligands to the inhibitors present in crystal structures suggests that the modification of the (S)-PMPDAP inhibitor, in which a 2-(phosphonomethoxy)propyl chain is attached at position 8 instead of position 9, might increase its binding affinity.



1. INTRODUCTION

Protein nucleoside phosphorylase (PNP) is an enzyme that catalyzes reversible conversion process (ribosylation and phosphorolysis) between nucleobases (purines) and their nucleosides. PNP plays an important role in nucleotide metabolism because it participates in a salvage metabolic pathway of nucleotide synthesis, which utilizes nucleobases and nucleosides available in the cell. This is an alternative pathway to a more common, but energetically more expensive, *de novo* synthesis process.¹ Such biochemical properties enable PNP to be utilized in pharmacological, medical, and practical processes. One of the negative consequences of the activity of the enzyme is phosphorolysis of nucleoside drugs, which can be attenuated by the application of suitable inhibitors.² On the other hand, phosphorylases may contribute to the activation of prodrugs that can be either nucleosides or nucleobases.³ Some of the PNP enzymes can be used in cancer gene therapy, in which cytotoxic nucleic acids are released as a result of nontoxic nucleoside phosphorolysis.⁴ It was shown that the deficiency or lack of PNP activity leads to dysfunction of T-cells and causes decreased cell immunity. PNP can also be used as an immunosuppressive drug for transplant rejection, drug for cancers causing overproduction of T-cells, and drug

for autoimmune diseases such as gout, rheumatoid arthritis, psoriasis, and multiple sclerosis.^{5–7} On the other hand, the enzymatic activity of PNP can be utilized for the synthesis of specific nucleosides, which stands as an alternative method for chemical synthesis.^{8–11}

A very interesting and important group of PNP substrates are purine analogues in which the carbon atom located at position 8 is replaced by nitrogen. 8-Azapurines contrary to canonical purines display measurable fluorescence emission and can be used for the investigation of kinetics of the binding process.¹² Fluorescence of 8-azaguanine is also pH-dependent and therefore can be used to probe the ionization states of nucleobases in structured RNAs.¹³ Recent experimental studies on 8-azapurines show that PNP can catalyze ribosylation in

Received: October 24, 2019

Published: January 16, 2020

various positions of the triazole ring, not necessary in canonical position 9.¹⁴

The reason for that diversity may emerge from differences in binding poses assumed by two investigated ligands. They may expose different nitrogens (7, 8, or 9) of a triazole ring to the channel leading to the binding site, which may cause ribosylation of ligands in these exposed positions. The complexity of this process is magnified by the presence of various tautomeric forms of ligands, as well as equilibrium of protonation states of amino acid residues present in the binding site. In this publication, various molecular modeling techniques are applied in an effort to answer the question: why are two investigated purine analogues ribosylated in different positions? The results are discussed in the context of available experimental data.

2. METHODS

In this publication, the binding of two analogues of purines, 2,6-diamino-8-azapurine (DaaPur) and 8-azaguanine (azaGua), to calf PNP protein is investigated. The methodology applied to the investigated system can be divided into three steps:

1. Determination of tautomeric equilibrium of the investigated ligands and protonation states of titrable residues of the protein.
2. Docking of selected lowest-energy tautomers to crystal structures of proteins, which were set up in the most probable protonation state.
3. Application of a more sophisticated methodology (molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) combined with normal mode analysis (NMA))^{15,16} to the selected binding poses obtained in the second step.

The methodology of each of these steps is described below.

2.1. Selection of Tautomeric Forms of Ligands.

Tautomeric equilibrium of the DaaPur molecule was determined in our recent publication.¹⁷ Here, the same methodology was applied to the second investigated guanine analogue, azaGua. The density functional theory BHandHLYP hybrid method¹⁸ combined with a triple-zeta valence cc-pVTZ basis set was applied.¹⁹ It was shown that the BHandHLYP/cc-pVTZ level of theory is suitable for the study of guanine derivatives such as 8-aza-iso-guanine¹⁷ and its methylated form,²⁰ and its efficiency is comparable to the composite Gaussian-3 method.²¹ The geometries of selected tautomeric forms of azaGua were optimized in the gas phase followed by optimization in water solution approximated by the IEF-PCM model.^{22,23} The vibrational analysis showed that all investigated tautomers reached a stationary point, and therefore thermochemistry calculations were performed. Finally, energy + zero point correction ($E + ZPE$), enthalpy ($H = E + k_B T$), entropy (S), and Gibbs free energy ($G = H - TS$) were determined. All quantum chemistry calculations were performed with Gaussian09 software.

The binding process was investigated only for tautomers with $\Delta G < 5$ kcal/mol (related to the lowest-energy tautomer) in water solution (see Table 1). It was assumed that populations of tautomers with higher Gibbs free energy can be neglected.

2.2. Protonation States and Preparation of Proteins.

Two crystal structures of calf PNP proteins were selected for further processing (PDB codes: 2A11 and 1LVU).^{24,25}

Table 1. Gibbs Free Energies and Populations of Various Tautomeric Forms of the DaaPur¹⁷ and azaGua Molecules in Water Solution^a

DaaPur			azaGua		
tautomer	ΔG (kcal/mol)	population (%)	tautomer	ΔG (kcal/mol)	population (%)
D1	13.7	0.0	A17	4.4	0.1
D3	6.6	0.0	A18	2.9	0.8
D7	7.7	0.0	A19	0.0	99.1
D8	3.9	0.2	A37	7.7	0.0
D9	0.0	99.8	A38	6.7	0.0
			A39	8.7	0.0
			A67c	12.8	0.0
			A67t	16.2	0.0
			A68c	12.3	0.0
			A68t	14.0	0.0
			A69c	7.9	0.0
			A69t	9.3	0.0

^aTautomeric forms with energies lower than the cutoff energy of 5 kcal/mol (with respect to the lowest-energy tautomer) are marked with bold face. Only these tautomers were considered in further investigations.

Coordinates of heavy atoms of both structures were downloaded from the RCSB protein data bank. The 2A11 and 1LVU structures were selected because of the structural and geometrical similarities of the ligands present in their binding sites to the azaGua and DaaPur molecules, respectively. The 1LVU structure contains coordinates of two trimers in the asymmetric unit, and the monomer D was selected because the number of missing residues was the lowest in this unit. The 2A11 molecule is represented in the form of a monomer (asymmetric unit), with several missing residues as well. Missing coordinates (residues: 1,253–256,285–289 for 1LVU and residues: 1,2,60–65,283–289 for 2A11) in both structures were determined by the application of comparative modeling using Modeller software.^{26–29} In this method, geometries of known sequences are reconstructed by the satisfaction of spatial restraints and optimization of the energy. For both molecules, missing fragments are either loop regions or short fragments of N- and C-termini; therefore, a dedicated Modeller tool for loop modeling and refinement was applied. Reconstructed monomers of the 1LVU and 2A11 molecules are shown in Figure 1. pK_a values of titratable residues were determined in experimental conditions¹⁴ of pH (6.6) using H+ + Web Server to determine their most probable protonation states.^{30–32}

Additionally, protonation states of all titrable residues were determined with 2.4 μ s long constant pH molecular dynamics simulation of the apo form of protein in explicit solvent. The 1LVU protein structure without a ligand was solvated in truncated octahedron filled with TIP3P water.³³ molecules. The initial protonation states were set to the standard values, and the structure was optimized with 1000 steepest descent steps followed by 4000 conjugate gradient steps. Harmonic restraints on backbone heavy atoms were applied. Then, the system was heated from 50 to 300 K during a 400 ps long NVT simulation, which was followed by a 4 ns long NPT equilibration. The simulation was run at pH = 6.6 with a 2 fs time step, and the titration of residues was performed every 100 steps. Each titration step was followed by 100 steps of water relaxation dynamics. The salt concentration for the titration process was

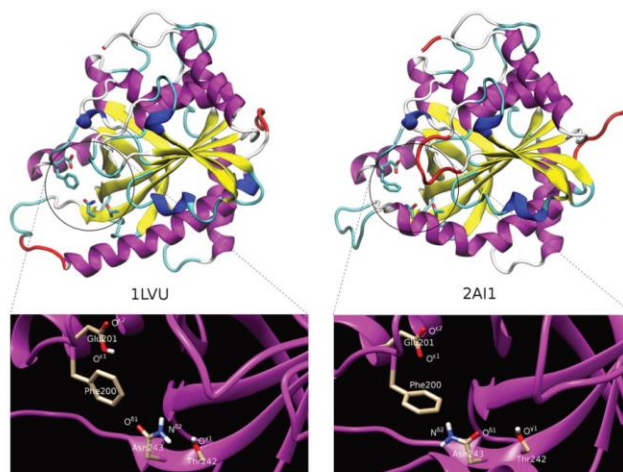


Figure 1. Tertiary crystal structures of calf PNP monomers, 1LVU (left) and 2A11 (right). Secondary structures are marked with magenta (α -helices), dark blue (3_{10} helices), yellow (β -sheets), and blue (turns). Fragments of proteins reconstructed with MODELLER tools are shown with a red ribbon. Glu-201, Asn-243, Thr-242, and Phe-200 binding site residues are visualized and zoomed on the bottom panels. χ_2 dihedral of Phe-200 in the 2A11 structure (bottom right panel) differs by 90° compared to the same dihedral in the 1LVU structure (bottom left panel) and also χ_2 dihedral of Asn-243 differs by 180° . Modified version of the 2A11 binding site, in which the χ_2 dihedral angle is changed by 180° , was also considered. This configuration of Asn-243 was found in the 2A13 crystal structure.

set to 0.1 M. The cutoff for electrostatic interactions was set to 8 Å, and long-range electrostatics was treated with the particle-mesh Ewald summation method.³⁴ The system was coupled to the heat and pressure reservoirs with the Langevin method with collision frequency set to 5.0 ps^{-1} . The protonation state of titrable residues was collected every 1000 steps. All simulations were performed with the AMBER16 molecular dynamics simulation package.³⁵ The uncertainties of pK_a 's were estimated by dividing the whole simulation into five 0.48 μs long intervals and calculating the errors from Student's t -distribution with a standard confidence level of 68%.

2.3. Docking. AutoDock Tools were used for the preparation of proteins and ligands for docking, i.e., definition of binding sites and positions and dimensions of the docking box. Binding poses and respective scoring functions of investigated complexes were determined with the AutoDock Vina molecular docking program.³⁶ All ligands and proteins were protonated before launching the docking procedure. The ligands were placed in the cuboid box of dimensions of $14 \text{ \AA} \times 12 \text{ \AA} \times 12 \text{ \AA}$ created in the binding pocket of the receptor. Based on 2A11 and 1LVU crystal structures, boxes included four residues crucial for the binding process: Phe-200, Glu-201, Thr-242, and Asn-243. For each complex (three tautomers of azaGua combined with 2A11 with two possible χ_2 rotamers Asn-243 (see Section 3.4) and two tautomeric forms of DaaPur in complex with 1LVU with protonated and deprotonated Glu-201), 10 binding poses were generated; therefore, a total of 60 binding poses for azaGua-2A11 and 40 poses for DaaPur-1LVU complexes were generated.

2.4. Application of MM-PBSA/NMA Methods to the Selected Binding Poses. The difference between the lowest and the highest scoring functions obtained for different binding poses does not exceed 1.2 kcal/mol (see Tables 2 and 3), and

Table 2. Scoring Functions of DaaPur Docked to 1LVU Receptor^{a,b}

pose	D8d	D8	D9d	D9
1	-5.9	-6.3	-5.8	-6.3
2	-5.8	-6.3	-5.8	-6.2
3	-5.7	-6.1	-5.7	-6.0
4	-5.7	-6.0	-5.6	-6.0
5	-5.6	-5.8	-5.5	-5.9
6	-5.6	-5.7	-5.2	-5.7
7	-5.3	-5.7	-5.2	-5.7
8	-5.2	-5.6	-5.1	-5.7
9	-5.2	-5.2	-5.1	-5.7

^aBinding poses (modes) qualified for the application of the MM-PBSA/NMA procedure are marked in bold. All MD-qualified modes have lowest scoring function except mode 3 for tautomer D9. ^bFor symbols of tautomers, see Figure 3. Symbol d marks systems with deprotonated Glu-201 binding site residue.

therefore a more sophisticated MM-PBSA/NMA method^{15,16} must be applied for the estimation of free energies of binding for selected binding poses obtained from the docking procedure. The most relevant criteria to sieve out appropriate binding poses were scoring function, number of established hydrogen bonds, proton clashes, and exposition of chemical groups, or atoms into the binding channel. The last criterion eliminates these binding poses, which expose the amino or keto group into the binding channel, which, by definition, prevents ribosylation. The MM-PBSA/NMA procedure is based on trajectories obtained from molecular dynamics simulations of protein–ligand complexes.

In the first step, molecular mechanics parameters of ligands in all considered tautomeric forms were determined. Bonding and Lennard-Jones parameters of ligands were obtained from

Table 3. Scoring Function of azaGua Docked to 2A11 Protein^{a,b}

pose	A17	A17m	A18	A18m	A19	A19m
1	-6.8	-6.8	-6.7	-6.6	-6.6	-7.0
2	-6.5	-6.6	-6.4	-6.6	-6.3	-6.4
3	-6.4	-6.1	-6.1	-6.5	-6.2	-6.2
4	-6.1	-6.1	-6.1	-6.4	-6.2	-6.2
5	-6.1	-6.1	-6.0	-6.3	-6.1	-6.1
6	-6.1	-6.1	-6.0	-6.0	-6.0	-6.1
7	-6.0	-6.1	-5.9	-6.0	-5.9	-6.0
8	-5.9	-6.0	-5.8	-5.8	-5.8	-5.9
9	-5.8	-6.0	-5.7	-5.7	-5.7	-5.8

^aBinding poses (modes) qualified for the MM-PBSA/NMA procedure are marked in bold. In contrast to 1LVU-DaaPur, roughly half of the selected modes are the most favored in respect to the scoring function results. ^bFor description of symbols of tautomers, see Figure 2 in the main text. Symbol m means that Asn-243 was reoriented in the binding site.

the general Amber force field (GAFF2),³⁷ and partial charges were determined by the application of the RESP procedure³⁸ to the electrostatic potentials generated around energy-optimized structures using quantum chemistry methods (HF/6-31G* level of theory in Gaussian09). The Antechamber tool³⁹ from the AMBER16 package was used for parametrization of ligands. Topology and coordinate files of ligands and complexes were prepared using the Leap program.⁴⁰ Partial charges of all ligands used in molecular dynamics simulations (D8, D9, A17, A18, and A19) are collected in Table S1 of the Supplementary Information.

For parametrization of proteins, the ff14SB force field was applied.⁴¹ The protonation states of all titratable residues were adjusted according to results obtained from H++ Web Server. Each complex was solvated in a truncated octahedron box filled with TIP3P water molecules.³³ The minimum distance from the surfaces of the box to the atoms of the complex was set to 12.5 Å. The following equilibration procedure was applied. Energies of complexes were minimized with 2000 steps of steepest descent followed by 3000 steps of conjugate gradient methods and then heated from 100 to 300 K with harmonic restraints (force constant 100 kcal/(mol Å²)) applied to all protein and ligand atoms. The equilibration procedure was divided into five 1 ns long steps performed in the NPT ensemble with a 0.5 fs time step. The first two steps were performed with all solute atoms restrained using harmonic restraints with force constants of 100 and 10 kcal/(mol Å²). The remaining three steps were performed with harmonic restraints of decreasing strength: 10, 1, and 0.1 kcal/(mol Å²) applied to the backbone heavy atoms, selected binding site residues (Phe-200, Glu-201, Thr-242, and Asn-243), and the ligand. These residues were indicated either in crystal structures or by the docking procedure as the crucial for coordination of ligands in the binding pocket of protein (see the bottom panels of Figure 1). The final production run was performed with a 2 fs time step in the NPT ensemble with temperature–pressure baths coupled with the Langevin algorithm with a target temperature of 300 K and a pressure of 1 atm. Long-range electrostatics was treated with the particle-mesh Ewald summation algorithm.³⁴ For all selected binding poses, 10 independent (with different initial velocities) 1 ns long production runs were carried out and the binding enthalpy $\Delta\Delta H_{\text{bind}}$ was determined using the MM-PBSA

method as implemented in the AMBER program suite.⁴² The MM-PBSA procedure was applied to 100 frames obtained from molecular dynamics (MD) trajectories with relative dielectric constants equal to 1 and 80 for interior and exterior, respectively, and an ionic strength of 0.1 mM. A protein–ligand entropy change upon binding ΔS values was calculated with the NMA method for each production run of selected binding poses using coordinates of the whole protein–ligand system. Twenty-five snapshots of each MD trajectory were used with 1000 minimization steps and the same ionic strength as for the MM-PBSA method. The last two thermodynamic parameters ΔH_{bind} and ΔS contribute to the binding Gibbs free energy: $\Delta G_{\text{bind}} = \Delta H_{\text{bind}} - T\Delta S$.

3. RESULTS AND DISCUSSION

3.1. Tautomeric Forms of Ligands. Gibbs free energies of various tautomers of DaaPur, both in the gas phase and in water solution, were determined in our previous publication.¹⁷ In this study, only the results obtained in water were utilized, and they are presented in Table 1. The dominant tautomer has proton at position 9, and the second lowest-energy tautomer is protonated at position 8 with $\Delta G \approx 4$ kcal/mol. In this publication, we set the energy cutoff for tautomeric forms to 5 kcal/mol, assuming that populations of molecules with higher free energy are negligible in water solution. Therefore, the docking procedure was applied only to D9 and D8 tautomeric forms of the DaaPur molecule.

The neutral amino form of azaGua contains two tautomeric protons; therefore, 12 tautomeric combinations are possible, as shown in Figure 2. Tautomeric picture is clarified by the following thermodynamic parameters: energy (ΔE), enthalpy (ΔH), free energy (ΔG), entropy component ($T\Delta S$), and free energy of solvation ($\Delta\Delta G$). All thermodynamic parameters computed for azaGua in the gas phase and in water solution are shown in Tables S2 and S3 of SI, respectively. For the purpose of this study, the most important thermodynamic

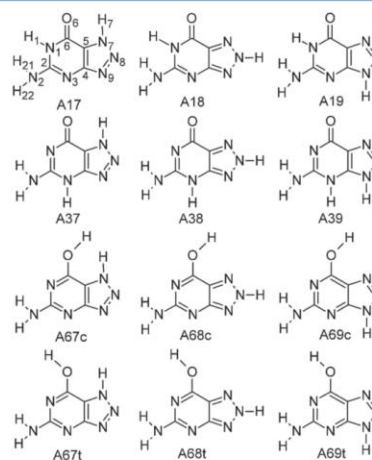


Figure 2. Tautomeric forms of 8-azaguanine. First and second rows present keto structures, while the third and fourth rows show cis and trans enol structures, respectively.

parameters are relative Gibbs free energies (ΔG) determined in water solution. These parameters for investigated tautomers and their respective populations are shown in Table 1.

3.1.1. Gas Phase. The relative Gibbs free energy of all keto tautomers protonated at position 1 (A19, A18, and A17) and two enol forms (A69c and A69t) is lower than 3 kcal/mol, and one can assume that their populations are significant (see Table S2). The most stable tautomer is A19 with small free energy favor ($\Delta G = 0.9$ kcal/mol) over A18—the next most preferred form, closely followed by the enol-cis A69c form (0.5 kcal/mol). This tautomeric order is reversed in comparison to Guanine,^{43,44} for which A17 is the most stable form, followed by A19 tautomer, although the free-energy difference is less than 0.5 kcal/mol. On the other hand, tautomers A39 and A67t are strongly disfavored, and the free energy difference between these two forms and A19 exceeds 14 kcal/mol. This phenomenon can be explained by the short distance between the proton located on the triazole ring and amino (A39) or enol (A67t) proton, which causes strong electrostatic repulsion. This effect also explains the tautomer balance shift toward the cis form observed for enol tautomers. Also, the negative correlation of the distance between protons of the triazole ring and the hydroxyl group and relative Gibbs free energy of the molecule can be observed for enol tautomers.

3.1.2. Aqueous Solution. Immersion of azaGua in water significantly destabilizes enol tautomeric forms (ΔG higher than 7 kcal/mol), which are abundant in the gas phase, but N(1)–H tautomers remain still preferred. The A19 form clearly prevails, followed by A18 and A17; however, free-energy differences between the last two forms and dominator increased by roughly 2 kcal/mol. Biologically, the most important position is N(9), which is the most probable place of ribosylation. Quantum calculations of Guanine tautomerism^{43,44} also revealed the highest stability of A19 tautomer and its undeniable domination among N(9)–H tautomers, but the second most favored Guanine tautomer N(1,7)–H is significantly more populated ($\Delta G < 1.0$ kcal/mol) than the corresponding azaGua tautomer ($\Delta G = 4.4$ kcal/mol).

The relative solvation effect is favorable only for N(3)–H forms ($\Delta\Delta G \leq -1$ kcal/mol). Despite the most favorable solvation free energy ($\Delta\Delta G = -9.5$ kcal/mol), the very unstable tautomer A39 in the gas phase does not become significantly populated in water solution ($\Delta G = 8.7$ kcal/mol). Water environment similarly affects the stability of most enol tautomers (except A67t), which is evidenced by the relatively low difference of $\Delta\Delta G$ (<2 kcal/mol). As in the gas phase, electrostatic repulsion determines cis–trans enol stability in favor of the latter one. Our results significantly differ from those obtained in the previous quantum chemical study of this molecule, in which energies were computed at the MP2/6-311++G(d,p) level using HF/6-31G(d) optimized structures.⁴⁵ In that study, tautomers were preliminarily sieved out using the semiempirical AM1 method, leaving only three tautomers A17, A19, and enol A69 for higher-level calculations. Our calculations show that A18 should also be considered, as it is the second most populated tautomer both in the gas phase ($\Delta G = 0.9$ kcal/mol) and in water solution ($\Delta G = 2.9$ kcal/mol). The previous study also shows that populations of A17 and A19 tautomers in the gas phase and in water solution are practically equal ($\Delta G = 0.22$ and -0.18 kcal/mol, respectively). This result cannot be confirmed by our calculations, which show a significant relative stabilization of A19 tautomer both in the gas phase ($\Delta G = 2.8$ kcal/mol) and

in water solution ($\Delta G = 4.4$ kcal/mol), but the relative stability of the A69 enol tautomers in the gas phase is practically equal in both studies ($\Delta G \cong 1.4$ kcal/mol). Our results also show that enol tautomers are much less stable in water solution ($\Delta G \cong 8.5$ kcal/mol) compared to the previous study ($\Delta G = 0.58$ kcal/mol). Following our assumption of 5 kcal/mol cutoff for the Gibbs free energy of tautomeric systems in water, only three tautomers—A19, A18, and A17—were docked to the target protein.

3.2. Protonation States of Titrable Residues. pK_a of the only titrable residue of the binding site (Glu-201) and pK_a 's of all remaining titrable residues of apo-ILVU protein, obtained with H++ Web Server, are shown in Table S4 of SI. The most important conclusion from theoretical titration is that pK_a of Glu-201 deviates significantly from experimentally determined pK_a of glutamic acid guest residue in a model alanine host pentapeptide in water solution (~ 4.25).⁴⁶ This effect is a consequence of electrostatic interactions with surrounding charges of the protein. The pK_a of Glu-201 calculated with H++ is 5.8, which is close to the experimental pH, and therefore both protonated and deprotonated (charged) forms of this residue must be considered in both docking and MM-PBSA simulation steps.

To confirm this important observation, constant pH explicit solvent molecular dynamics simulation of ILVU protein in the apo form was performed. The calculated pK_a 's (Table S4) are in qualitative agreement with the results obtained with H++ in a sense that they point to the same most probable ionization states for all residues except His-64, which should be deprotonated with $pK_a = 5.78$ (H++ ~ 7.1). It is not clear what is the most probable ionization state for His-175 for which pH-offset is equal to zero (H++) and 0.1 ± 0.3 (constant pH MD), but as the distance of this residue from binding site is approximately 24 Å, it was assumed that its influence on the binding affinity is small and the deprotonated form of this residue was picked for MM-PBSA MD simulations. However, the most important result obtained from constant pH simulation is that pK_a of Glu-201 equal to 5.85 ± 0.11 almost exactly matches the one obtained with H++ Web Server (5.8), which means that the population of the protonated form at pH = 6.6 is about 10% and it cannot be neglected. Therefore, the most probable protonation states of all titrable residues obtained from H++ were assumed (see Table S4) with the exception of a very important Glu-201 residue, which was considered both in its protonated and deprotonated forms. It should be stressed here that constant pH simulations show that there are four His (64, 86, 230, and 257) and three more Glu (52, 89, and 109) residues with populations of protonated forms exceeding 10%, but their distance from the binding site is at least 8 Å and therefore their influence on the binding affinity of ligands was neglected in the first approximation.

Another interesting conclusion that can be drawn from a comparison of H++ and constant pH MD data is that pK_a 's obtained with the former method lead to significantly higher populations of negatively charged residues ($\overline{pK_a}_{Asp} = 3.0$ and $\overline{pK_a}_{Glu} = 3.6$) than pK_a 's obtained from constant pH MD ($\overline{pK_a}_{Asp} = 4.3$ and $\overline{pK_a}_{Glu} = 5.2$). The same effect is also observed for titrable positively charged residue with $\overline{pK_a}_{Lys} = 10.8$ (H++) and $\overline{pK_a}_{Lys} = 8.6$ (constant pH MD). The effect is most probably due to the conformational

averaging of constant pH MD simulation, which is absent in “single point” H++ computations.

3.3. Binding of DaaPur to Calf PNP. Two most populated tautomeric forms of DaaPur, namely, D8 and D9 (see Figure 3), based on quantum calculations,¹⁷ were selected

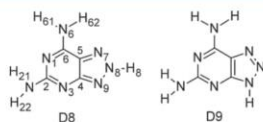


Figure 3. DaaPur most populated tautomeric structures. In contrast to azaGua, only one tautomeric proton is located on the entire structure.

to examine their binding affinity to calf PNP enzyme (pdbcode: 1LVU). The results of 1LVU titration in experimental pH showed that Glu-201 can exist in either the protonated or deprotonated form, and therefore both selected DaaPur tautomers were docked to the binding pocket with either the protonated or deprotonated Glu-201 residue. Therefore, there are four possible ligand–receptor systems, as can be seen in Table 2. Eleven out of 60 binding poses were qualified for further molecular dynamics simulations, and their scoring functions are shown with a bold face type in Table 2.

Most of binding systems are represented by three binding poses; the only exception is the deprotonated 1LVU-D9 system for which only two representatives were qualified. For systems with D8 and D9 tautomers, three and two binding poses with lowest scoring functions were qualified for MM-PBSA. Additionally, the fourth binding pose for the D9 ligand–receptor system was qualified. The results of docking computations clearly showed various possible orientations of the ligand in the binding pocket, which leads to different expositions of the triazole ring into the binding channel, as can be seen in Figure 3 and Figures S1 and S2 (S1). Results of docking also showed that the exposed position of the ligand can be either protonated or deprotonated.

The results of the application of the MM-PBSA/NMA method to the selected binding poses of the 1LVU-DaaPur ligand–receptor system are collected in Table 4. The results of

binding free energy clearly show strong favor of binding poses with the protonated Glu-201 residue ($\Delta\Delta G_{\text{bind}} = 2.5$ kcal/mol - difference between the lowest deprotonated—D8dM3 - and the highest protonated—D8M2 - Glu-201 modes). All binding poses with protonated Glu-201 have lower free energies than those with deprotonated residue, as can be seen in Table 4. This important result shows that the DaaPur molecule must bind to the protein with the protonated Glu-201 residue. The most favored binding poses are D8M1 and D9M2 with the identical hydrogen-bond interaction pattern, as shown in Figure 4, and similar binding free energy with $\Delta\Delta G_{\text{bind}} = 0.3$ kcal/mol. Both binding poses are stabilized by the following hydrogen bonds between Glu-201 and DaaPur: O^{δ2}–N6, O^{δ1}–N1, and Asn–243 and DaaPur: O^{δ1}–N2 and N^{δ2}–N3. Figure 4 shows the solvent-accessible surface area of the binding pocket with ligands for D8M1 and D9M2 binding poses. It can be clearly seen that for both ligands, position N8 is the one that is most exposed to the binding channel and therefore it is the most probable place of ribosylation. This result is partially convergent with experimental data, which show a balance between N8 and N7 ribosylated forms.¹⁴ Although the next three lowest free-energy binding poses expose either N9 or N3 nitrogens, the sixth binding pose, which exposes N7 nitrogen, cannot be completely ruled out. A relatively low position of this binding pose in free-energy ranking is caused by the highest entropy drop among all investigated binding poses. Consideration of only the enthalpic part of free energy obtained from MM-PBSA, as suggested by others,⁴⁷ places N7-exposing binding pose in the fourth position, just 0.1 kcal/mol behind the third one D9M1 (see Table 4).

The next two thermodynamically most stable binding poses of D8 (D8M3 and D8M2 shown in Figure S1) and D9 (D9M1 and D9M4 shown in Figure S2) have significantly higher $\Delta\Delta G_{\text{bind}}$ (9.8 kcal/mol at least). D8M2 and D9M1 pairs expose N9 nitrogen. D8M3 and D9M4 binding poses expose N3 nitrogen of pyrimidine ring. These binding poses, despite their energetical disfavor, also decrease the cavity for the ribosylation process caused by the presence of an amino group.

3.4. Binding of azaGua to Calf PNP. The application of a 5 kcal/mol free-energy cutoff for azaGua tautomers in water leaves only three forms protonated at position 1 and positions

Table 4. Enthalpy (MM-PBSA), Entropic Contribution ($T = 300$ K, NMA), and Gibbs Free Energy (kcal/mol) of the Binding Process in the DaaPur–Calf PNP (1LVU) System Obtained with the MM-PBSA/NMA Method^{a,b}

	binding pose	ΔH_{bind}	$-T\Delta S$	ΔG_{bind}	$\Delta\Delta H_{\text{bind}}$	$\Delta\Delta G_{\text{bind}}$	exposed nitrogen
1	D8M1	-11.4 ± 1.6	17.9 ± 1.0	6.6 ± 1.9	0	0	8
2	D9M2	-10.5 ± 1.5	17.5 ± 0.8	6.9 ± 1.7	0.9	0.3	8
3	D9M1	-7.5 ± 1.7	17.3 ± 0.8	9.8 ± 1.9	3.9	3.2	9
4	D8M3	-5.1 ± 1.7	16.7 ± 1.3	11.5 ± 2.1	6.3	4.9	3
5	D9M4	-5.4 ± 1.5	16.9 ± 0.8	11.5 ± 1.7	6.0	4.9	3
6	D8M2	-7.4 ± 0.9	19.2 ± 1.5	11.9 ± 1.7	4.0	5.3	7
7	D8dM3	-1.6 ± 1.8	16.0 ± 0.9	14.4 ± 2.0	9.8	7.8	7
8	D8dM2	-1.5 ± 1.6	16.2 ± 1.2	14.7 ± 2.0	9.9	8.1	8
9	D9dM2	0.2 ± 1.8	15.0 ± 1.4	15.2 ± 2.3	11.6	8.6	7
10	D9dM1	-0.6 ± 2.0	16.0 ± 0.7	15.4 ± 2.1	10.8	8.8	8
11	D8dM1	1.0 ± 1.6	18.3 ± 1.0	19.3 ± 1.9	12.4	12.7	3

^a $\Delta H_{\text{bind}} - T\Delta S$, and ΔG_{bind} are averages of 10 independent runs. ^bThe relative enthalpy and Gibbs free energy with respect to the lowest-energy binding pose are shown in the sixth and seventh columns, respectively. The number of exposed nitrogens is shown in the last column. Two lowest free-energy binding poses are shown in Figure 4; other binding poses are shown in Figure S1 (D8 tautomer) and Figure S2 (D9 tautomer). Binding poses with D9 tautomer are marked with a bold face type.

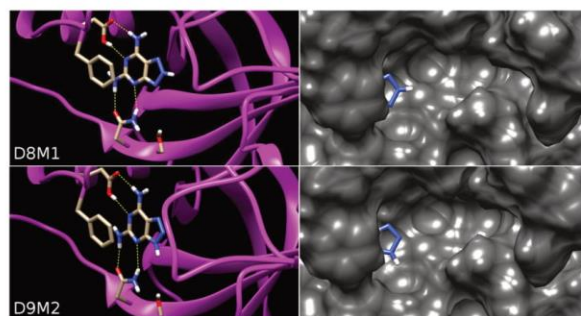


Figure 4. Interaction map and binding site for two DaaPur most favored binding poses. Glu-201 and Asn-243 are both donor and acceptor of hydrogen bonds, and Phe-200 is engaged in stacking contact. Binding cavity is visualized by the solvent-accessible surface area. Letter code: D8M1 means the first binding mode (scoring function order) of 1LVU–D8 complex.

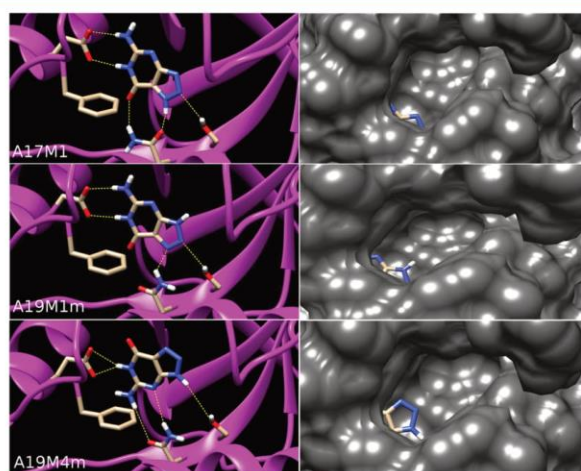


Figure 5. Visualization of the protein–ligand interaction and binding cavity for the most preferred binding poses of azaGua. Glu-201, Asn-243, and Thr-242 are involved in hydrogen bonds, while Phe-200 is responsible for stacking interactions. The binding site is shown as the solvent-accessible surface area. Letter code: A19mM4 means the fourth binding mode (M4) of the 1LVU–A19 complex with Asn-243 modified (m).

7–9. As for the 1LVU–DaaPur system, residues Phe-200, Glu-201, and Asn-243 are crucial for the binding process of azaGua to 2A11, although docking results also revealed that Thr-242 residue can form hydrogen bonds with ligands. Titration results for 2A11, performed with H⁺⁺, showed identical protonation balance of Glu-201 as for 1LVU. However, docking to the protein with protonated Glu-201 does not seem to be necessary due to a possible sterical clash of the proton located at position 1 on selected azaGua ligands and the proton of Glu-201. Nevertheless, docking of azaGua to 2A11 protein with protonated Glu-201 was performed, but none of the resulting structures were qualified for the application of the MM-PBSA/NMA method. This result leads to an important conclusion that binding of most abundant azaGua tautomers requires PNP protein with the deprotonated Glu-201 residue.

There are three similar crystal structures of calf PNP protein in the RCSB database with different multisubstrate inhibitors—2A11 (guanosine-2',3'-O-ethylidene phosphonate inhibitor), 2A12 (9-deazainosine-2',3'-O-ethylidene phosphonate inhibitor), and 2A13 (guanosine-2',3'-O-methylidene phosphonate inhibitor).²⁴ The purine parts of multisubstrate analogue in 2A11 and 2A13 are guanines, and therefore both structures can be used for docking of the azaGua molecule. The side chain of one of the ligand-coordinating residues Asn-243 has a different conformation in the 2A13 structure compared to that of 2A11. Their χ_2 dihedral angles differ by 180°, and therefore both possible rotamers were considered in the docking procedure. Systems with an alternative rotamer of Asn-243 are marked with the letter m, e.g., A19mM1 describes binding mode 1 (M1) with the azaGua molecule protonated at positions 1 and 9 (A19) and modified Asn-243 rotamer from the 2A13 crystal structure (m), as shown in Figure 5. The

Table 5. Enthalpy (MM-PBSA), Entropic Contribution ($T = 300$ K, NMA), and Gibbs Free Energy (kcal/mol) of the Binding Process in the azaGua–Calf PNP (2A11) System Obtained with the MM-PBSA/NMA Method^{a,b}

	binding pose	ΔH_{bind}^0	$-T\Delta S$	ΔG_{bind}	$\Delta\Delta H_{\text{bind}}$	$\Delta\Delta G_{\text{bind}}$	exposed nitrogen
1	A17M1	-14.6 ± 2.0	18.6 ± 1.3	4.0 ± 2.4	0	0	9
2	A19mM4	-10.5 ± 1.4	16.6 ± 1.4	6.1 ± 2.0	4.1	2.1	8
3	A19mM1	-9.3 ± 1.4	17.2 ± 1.0	7.7 ± 1.7	5.3	3.7	9
4	A18mM6	-8.4 ± 1.4	17.5 ± 1.5	9.1 ± 2.1	6.2	5.1	9
5	A19M1	-7.6 ± 1.6	17.1 ± 0.8	9.5 ± 1.8	7.0	5.5	9
6	A18M5	-6.6 ± 2.9	16.2 ± 1.4	9.7 ± 3.2	8.0	5.7	3
7	A18mM7	-8.2 ± 2.6	18.0 ± 1.5	9.8 ± 3.0	6.4	5.8	8
8	A18M4	-7.2 ± 1.0	17.7 ± 1.4	10.5 ± 1.7	7.4	6.5	9
9	A18mM4	-6.1 ± 1.2	17.5 ± 1.9	11.4 ± 2.2	8.5	7.4	8
10	A17mM2	-4.8 ± 1.7	16.6 ± 2.0	11.8 ± 2.6	9.8	7.8	7
11	A19M2	-4.3 ± 3.9	16.2 ± 1.1	12.0 ± 4.1	10.3	8.0	7
12	A18mM1	-5.9 ± 2.6	18.0 ± 0.8	12.1 ± 2.7	8.7	8.1	9
13	A17M2	-5.9 ± 2.5	18.1 ± 1.3	12.2 ± 2.8	8.7	8.2	3
14	A17mM1	2.5 ± 5.9	16.6 ± 2.3	19.1 ± 6.3	17.1	15.1	7
15	A18mM3	3.7 ± 3.1	16.4 ± 1.2	20.0 ± 3.3	18.3	16	7

^a ΔH_{bind}^0 , $-T\Delta S$, and ΔG_{bind} parameters are averages of 10 independent runs. ^bThe relative enthalpy and Gibbs free energy with respect to the lowest-energy binding pose are shown in the sixth and seventh columns, respectively. The number of exposed nitrogens is shown in the last column. Three lowest free-energy binding poses are shown in Figure 5. Other binding poses are shown in Figure S3 (A17 tautomer), Figure S4 (A18 tautomer with modified rotamer of Asn-243 residue), and Figure S5 (A18 tautomer with unmodified Asn-243). Binding poses with A17, A18, and A19 are marked with normal, bold face, and italic type, respectively.

results of docking of azaGua showed 15 reasonable binding poses (see Table 3, Figure 5, and Figures S3–S6) with a predominant amount of A18m binding poses. Interestingly, for a system with unmodified Asn-243 and A18 tautomers, only the fourth and fifth lowest scoring function poses were selected for the MM-PBSA/NMA procedure. In contrast to DaaPur, only a half of binding systems (2A11—A17, A17m, A19) are represented by only the two lowest scoring function modes. Nevertheless, the lowest scoring function binding poses were qualified for MD simulations for all but one investigated system (A18). Docking results showed the possibility of exposition of all triazole ring nitrogens into the binding channel, regardless of their protonation state (see Figures 5 and S3–S6).

All results of the application of the MM-PBSA/NMA method to the qualified binding poses for the 2A11–azaGua system are shown in Table 5. The lowest binding free energy belongs to the A17M1 binding pose ($\Delta G_{\text{bind}} = 4.0$ kcal/mol; see Figure 5 and Table 5) that exhibits position 9 of the triazole ring into the binding channel, confirming the results of the experiment,¹⁴ in which the ability of azaGua to form canonical N9 ribosides is concluded. The contribution of the binding enthalpy ΔH_{bind}^0 singles out A17M1 with more than 4 kcal/mol predominance over modified A19mM4 (Table 5). The next most stable mode, A19mM4, is clearly behind the leader (ΔG_{bind} about 2 kcal/mol higher than A17M1), although the entropy term favors the latter one by roughly 2 kcal/mol. A19mM4 exposes position 8, which suggests that noncanonical alternative ribosylation is also possible. Thermodynamic calculations reveal one more possibly populated azaGua mode, A19mM1 ($\Delta G_{\text{bind}} = 7.7$ kcal/mol). For this mode, position 9 of the triazole ring is exposed (see Figure 5). On the other hand, modes exposing position 7 are very unlikely due to a high value of ΔG_{bind} (11.8 kcal/mol at least), and it also does not converge with experimental data, which excludes the possibility of ribosylation at position 7.¹⁴

3.5. Free Energies of Families of Binding Poses. Gibbs free energies of binding for both ligands are positive because

the positive entropic term exceeds the negative enthalpic term. This result is counterintuitive, especially for the azaGua ligand for which the experimentally determined dissociation constant $k_{\text{d}} = 90 \mu\text{M}$ ⁴⁸ converts to the free energy of binding $\Delta G_{\text{exp}} \approx -5.5$ kcal/mol (dissociation constant for DaaPur is not available). This mismatch is caused by a relatively large positive nonpolar solvation energy term generated by the model in which it is a sum of dispersion and cavity formation terms.⁴⁹ The alternative, older model of nonpolar interactions, based on the solvent-accessible surface area (SASA),⁵⁰ generates (on average) more than 10 kcal/mol lower nonpolar energy terms. Although the latter model of nonpolar interactions provides results that are closer to the absolute value of binding free energy for azaGua, we decided to use the former one because it was shown that it leads to better agreement with the experimental data of relative free energies for protein–ligand systems⁵¹ and significantly improves correlation between the cavity free energy and SASA (or molecular volume enclosed by SASA) for cyclic organic molecules.⁵² In this publication, we are interested in the relative binding free energies of various binding poses, and therefore the former, “modern” model of nonpolar interactions seems to be more suitable for that purpose.

To answer the question of what is the most probable place of exposition of investigated ligands, binding poses were divided into families with respect to their binding site exposition nitrogen. There are four “exposition” families for each ligand with the following exposed nitrogens: N3, N7, N8, and N9. The members of each exposition family are shown in Table 4 (DaaPur) and Table 5 (azaGua). It can be shown that the Gibbs free energy of each family can be determined from the following equation

$$\Delta G = -k_{\text{B}}T \ln \left[\sum_{i,j} f_{\text{P}_i} f_{\text{L}_j} \exp \left(-\frac{\Delta G_{ij}}{k_{\text{B}}T} \right) \right] \quad (1)$$

where f_{P_i} is the population of the i th ionization state of protein, f_{L_j} is the population of the j th tautomeric state of a ligand, ΔG_{ij}

is the binding free energy of ligand L_i to the protein P , k_B is the Boltzmann constant, and $T = 298$ K is the absolute temperature. Equation 1 includes corrections of free energies, which comes from populations of ionization states of proteins and populations of tautomeric states of ligands. In the first approximation, only varying protonation of residues of the binding site was considered. Therefore, two possible ionization states of the protein with protonated ($f_p = 0.1$) and deprotonated ($f_p = 0.9$) Glu-201 residues were applied to eq 1. Populations of tautomeric states used in eq 1 for DaaPur and azaGua ligands are shown in Table 1. The application of eq 1 to the above-mentioned populations of ligands and proteins together with the binding free energies of various binding poses obtained from Table 4 (DaaPur) and Table 5 (azaGua) leads to Gibbs free energies of each exposition family shown in Table 6.

Table 6. Ligand and Protein Population-Corrected Gibbs Free Energy and Enthalpy Differences (kcal/mol) of Four Exposition Families of Two Investigated Ligands^a

exposed nitrogen	DaaPur		azaGua	
	$\Delta\Delta H$	$\Delta\Delta G$	$\Delta\Delta H$	$\Delta\Delta G$
N3	5.10	4.60	6.82	6.47
N7	6.79	6.96	6.27	5.90
N8	0.00	0.00	0.07	0.00
N9	3.01	2.90	0.00	1.34

^aGibbs free energies and enthalpies were computed according to eq 1. The differences were calculated with respect to the lowest free-energy (enthalpy) family.

The population-corrected analysis shows that the most probable exposition site of the DaaPur molecule is N8, but the probability of exposition of N7 is low. This result is in partial agreement with the experimental data, which predicts ribosylation at positions N8 and N7. The Gibbs free energy obtained with the same analysis for azaGua predicts the highest stability of N8-exposed binding poses, but experimentally predicted exposition of N9 is only 1.3 kcal/mol away from N8, certainly well within an error of the MM-PBSA method. Moreover, the enthalpic term suggests that the exposition of nitrogen N9 is the most probable one. Therefore, our data cannot unambiguously select the most probable place of exposition of the azaGua ligand, which can be either N8 or N9 nitrogen.

3.6. Comparison of Computed Binding Poses to the Crystal Structures. Both 2A11 (2A13) and 1LVU crystal structures were solved with multisubstrate analogues bound in their active site, guanosine-2',3'-O-ethylidene phosphonate (guanosine-2',3'-O-methylidene phosphonate) and 2,6-diamino-(S)-9-[2-(phosphonomethoxy)propyl]purine, respectively. The purine part of multisubstrate analogues bound to crystal structures is similar to purine analogues considered in this publication; therefore, our analysis should also include a comparison of binding poses of ligands present in crystal structures to those determined by our procedure.

The lowest free-energy binding pose of azaGua bound to PNP protein is very similar to the binding pose of the purine part of the multisubstrate analogue in the 2A11 crystal structure,²⁴ as can be seen in Figure 6a. This is an expected result because the experimentally determined place of ribosylation of azaGua¹⁴ is the same as the place of attachment of the sugar-phosphate moiety (2',3'-O-ethylidene phospho-

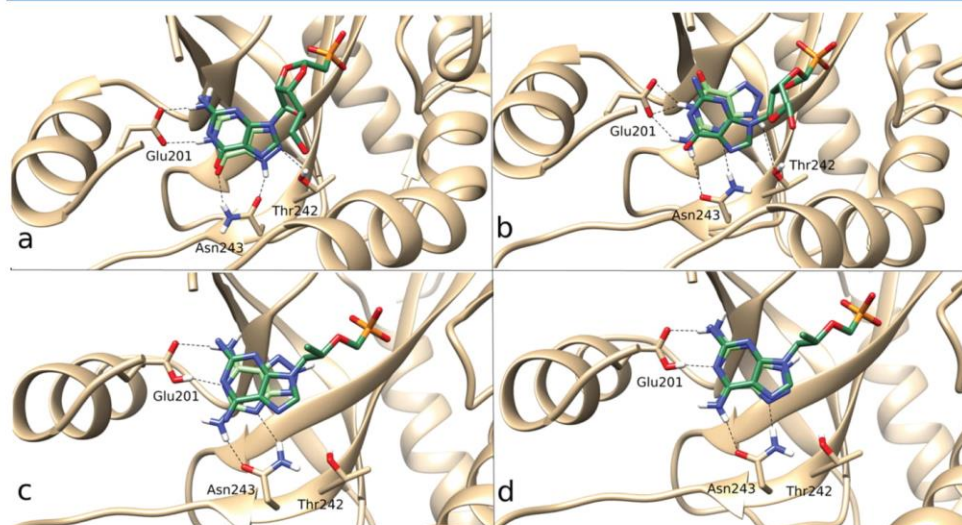


Figure 6. Comparison of the selected binding poses of DaaPur and azaGua with binding poses of original ligands bound to the crystal structures. (a) A17M1, the lowest free-energy binding pose of azaGua, (b) A19mM4, the lowest free-energy binding pose with the modified rotamer of Asn-243, (c) D8M1, the lowest free-energy binding pose of DaaPur, (d) D9M1, the third lowest free-energy binding pose of DaaPur. Carbon atoms of multisubstrate analogues and ligands considered in this publication are shown in dark green and light green colors, respectively.

nate) in the crystal structures (position 9). This result also suggests that exposition of N9 should rather emerge as a winner from population-corrected exposition family analysis performed in the previous section.

A comparison of the binding poses of DaaPur and respective multisubstrate analogue 2,6-diamino-(S)-9-[2-(phosphonomethoxy)propyl]-purine ((S)-PMPDAP) present in the binding pocket of the 1LVU structure seems to be much more interesting. Two lowest free-energy binding poses of DaaPur differ significantly from binding poses of the base part of (S)-PMPDAP present in the binding pocket of the 1LVU crystal structure,²⁵ as can be seen in Figure 6c. The multisubstrate ligand of 1LVU is "ribosylated" at position 9, and interactions of the 2-(phosphonomethoxy)propyl tail might force the ligand to assume a binding pose in which position number 9 is "exposed". The orientation of the ligand in the binding pocket of the crystal structure is similar to the third lowest free-energy binding pose, D9M1, which expose N9 nitrogen to the binding channel, as can be seen in Figure 6d. Assuming that substitution of carbon by nitrogen at position 8 does not significantly alter the coordination of the ligand, one can conclude that the coordination of the base fragment of ((S)-PMPDAP) is different from the coordination of 2,6-diaminopurine. The most probable binding poses obtained from our calculations (D8M1 and D9M1) prefer the exposition of N8 nitrogen, and therefore the binding of the multisubstrate analogue with the 2-(phosphonomethoxy)propyl tail attached to 2,6-diaminopurine at position 8 (namely, 2,6-diamino-(S)-8-[2-(phosphonomethoxy)propyl]-purine) might be stronger than the binding of the original ligand.

A similar conclusion can also be drawn for the guanosine-2',3'-O-methylidene phosphonate ligand bound to calf PNP. This ligand binds to the active site of PNP with a modified Asn-243 rotamer, as can be seen in the 2A13 crystal structure. Figure 6b shows the ligand from the 2A13 crystal structure overlapped with the lowest free-energy binding pose of azaGua. The difference between the two binding poses is clearly visible. Exposition of nitrogen N8 (rather than N9) also suggests that guanine with 2',3'-O-methylidene phosphonate moiety attached at position 8 might have higher binding affinity than the original multisubstrate analogue.

Another important consequence of this study is related to tautomeric equilibrium of ligand and protonation states of the crucial Glu-201 residue. It was suggested that 2,6-diaminopurine must be protonated at position 1 to make favorable hydrogen-bond contacts with the Glu-201 residue.²⁵ Our previous study¹⁷ shows that protonation of position 1 of the DaaPur molecule in water is very improbable ($\Delta G \approx 12$ kcal/mol with respect to the lowest free-energy D9 tautomer). This result combined with the results of computational titration of Glu-201 shows that DaaPur tautomer deprotonated at position 1 makes a hydrogen bond with the protonated Glu-201 residue.

4. CONCLUSIONS

The goal of this publication was the explanation of experimentally observed ribosylation of two purine analogues: 2,6-diamino-8-azapurine (noncanonical) and 8-azaguanine (canonical) by calf PNP protein. We postulated the hypothesis that the place of ribosylation depends on exposition of the triazole ring of ligands to the channel leading to the binding site, where α -D-ribose-1-phosphate can approach its target.

Consequently, the place of ribosylation depends on the most probable binding pose assumed by the ligand in the binding pocket, i.e., the binding pose with the lowest Gibbs free energy of binding. The hypothesis was verified by the application of a combination of quantum chemistry, docking, and molecular dynamics protocols (MM-PBSA and NMA methods), which pointed out the lowest free-energy binding poses for both investigated ligands.

Before the application of the docking procedure, the most probable tautomeric states of ligands were determined. Tautomeric equilibrium of the DaaPur molecule was determined in our previous publication.¹⁷ It was shown that the population of tautomeric forms protonated at position 1 is very low for this molecule with $\Delta G \approx 12$ kcal/mol. The results of the application of the same methodology to the azaGua molecule showed that all three most populated tautomeric forms in water solution are protonated at position 1 and differ in protonation of the triazole ring, which can be protonated at all three available positions: 9 (lowest free energy), 8 ($\Delta G = 2.9$ kcal/mol), and 7 ($\Delta G = 4.4$ kcal/mol higher). Only significantly populated tautomers ($\Delta G < 5$ kcal/mol) were picked for the docking procedure, i.e., D8 and D9 tautomers of DaaPur and A19, A18, and A19 tautomers of azaGua.

Two most probable binding poses of the DaaPur molecule inside the calf PNP binding pocket expose N8 nitrogen, partially in agreement with the experimental data, which shows a 1:1 ratio of products ribosylated at positions N7 and N8.¹⁴ Although the first binding pose, which exposes N7 nitrogen, is only the sixth lowest free-energy pose, its presence in vitro cannot be completely ruled out. The low position of this binding pose in the ranking is caused by its highest entropy drop among all investigated binding poses. If we consider only enthalpic contribution, as was suggested by some authors for the single trajectory MM-PBSA method,⁴⁷ this binding pose is the fourth lowest only 0.1 kcal/mol behind the third-placed one. For the other investigated ligand, 8-azaguanine, the most probable binding pose exposes N9 nitrogen, in agreement with experimental data.¹⁴ Although the second binding pose exposes N8 nitrogen, the next three binding poses again expose the nitrogen pointed out by the experimental data (N9). Analysis of Gibbs free energies and enthalpies of exposition families points to the exposition of N8 for DaaPur and either N8 or N9 for azaGua.

The other important result of this study is related to the protonation of crucial amino acid residue present in the binding pocket, Glu-201. Both theoretical titration with H++ Web Server and constant pH simulation of apo form of protein in explicit solvent showed that pK_a of its side chain is significantly higher than the one determined experimentally for model systems,⁴⁶ and therefore the population of protonated forms is not negligible and must be considered in the docking procedure.

For a long time, 6-oxo-substituted purines had been considered as a necessity for the specificity of trimeric PNP,¹ but it was later shown that the 6-amino purine analogue can also bind to the active site of this protein.²⁵ Analysis of populations of tautomers of the azaGua and DaaPur molecules, combined with theoretical titration of protein and docking/MM-PBSA/NMA analysis, shows that target proteins for 6-oxo- and 6-amino-substituted purine analogues are not identical. Our results suggest that the azaGua molecule (6-oxo substituent) binds to the protein with deprotonated Glu-201 residue, but DaaPur (6-amino substituent) binds to the

protein with the protonated Glu-201, which according to our theoretical titration data should also be significantly populated in water solution.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.9b00985>.

Partial charges of atoms of tautomers used in the study, all thermodynamic parameters of 8-azaguanine computed with quantum chemistry methods, results of computational titration of 1LVU protein in apo form, and high free-energy binding poses of the DaaPur and azaGua molecules docked to calf PNP protein (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Maciej Maciejczyk – Department of Physics and Biophysics, Faculty of Food Science, University of Warmia and Mazury in Olsztyn 10-719 Olsztyn, Poland; orcid.org/0000-0002-6138-7486; Email: maciej.maciejczyk@uwm.edu.pl

Author

Maciej Pyrka – Department of Physics and Biophysics, Faculty of Food Science, University of Warmia and Mazury in Olsztyn 10-719 Olsztyn, Poland

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jcim.9b00985>

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

This research was supported by the department internal Grant no. 17.610.011-110.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Some calculations were performed in the Wrocław Centre for Networking and Supercomputing (<http://wcss.pl>, grant number 349).

■ ABBREVIATIONS

PNP, purine nucleoside phosphorylase; azaGua, 8-azaguanine; DaaPur, 2,6-diamino-8-azapurine; (S)-PMPDAP, 2,6-diamino-(S)-9-[2-(phosphonomethoxy)propyl]-purine; MM-PBSA, molecular mechanics Poisson–Boltzmann surface area; NMA, normal mode analysis

■ REFERENCES

- (1) Bzowska, A.; Kulikowska, E.; Shugar, D. Purine Nucleoside Phosphorylase: Properties, Functions, and Clinical Aspects. *Pharmacol. Ther.* **2000**, *88*, 349–425.
- (2) Stoekler, J. D. Purine Nucleoside Phosphorylase: A Target for Chemotherapy. In *Developments in Cancer Chemotherapy*; Glazer, R.L., Eds.; CRC Press, Inc.: Boca Raton, FL, 1984; pp 35–60.
- (3) Mader, R. M.; Sieder, A. E.; Braun, J.; Rizovski, B.; Kalipcian, M.; Mueller, M. W.; Jakesz, R.; Rainer, H.; Steger, G. G. Transcription and Activity of 5-Fluorouracil Converting Enzymes in Fluoropyr-

imidine Resistance in Colon Cancer in Vitro. *Biochem. Pharmacol.* **1997**, *54*, 1233–1242.

- (4) Secrist, J. A., III; Parker, W. B.; Allan, P. W.; Bennett, L. L., Jr.; Waud, W. R.; Truss, J. W.; Fowler, A. T.; Montgomery, J. A.; Ealick, S. E.; Wells, A. H.; Gillespie, G. Y.; Gadi, V. K.; Sorscher, E. J. Gene Therapy of Cancer: Activation of Nucleoside Prodrugs with *E. coli* Purine Nucleoside Phosphorylase. *Nucleosides Nucleotides* **1999**, *18*, 745–757.

- (5) Ravandi, F.; Gandhi, V. Novel Purine Nucleoside Analogues for T-Cell-Lineage Acute Lymphoblastic Leukaemia and Lymphoma. *Expert Opin. Invest. Drugs* **2006**, *15*, 1601–1613.

- (6) Robak, T.; Lech-Maranda, E.; Korycka, A.; Robak, E. Purine Nucleoside Analogs as Immunosuppressive and Antineoplastic Agents: Mechanism of Action and Clinical Activity. *Curr. Med. Chem.* **2006**, *13*, 3165–3189.

- (7) Ho, M. C.; Shi, W.; Rinaldo-Matthis, A.; Tyler, P. C.; Evans, G. B.; Clinch, K.; Almo, S. C.; Schramm, V. L. Four Generations of Transition-State Analogues for Human Purinenucleoside Phosphorylase. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 4805–4812.

- (8) Mikhailopulo, I. A. Biotechnology of Nucleic Acid Constituents—State of the Art and Perspectives. *Curr. Org. Chem.* **2007**, *11*, 317–335.

- (9) Mikhailopulo, I. A.; Miroshnikov, A. I. Biologically Important Nucleosides: Modern Trends in Biotechnology and Application. *Mendeleev Commun.* **2011**, *21*, 57–68.

- (10) Stepchenko, V. A.; Seela, F.; Esipov, R. S.; Miroshnikov, A. I.; Sokolov, Y. A.; Mikhailopulo, I. A. Enzymatic Synthesis of 2-Deoxy- β -d-Ribonucleosides of 8-Azapurines and 8-Aza-7-Deazapurines. *Synlett* **2012**, *23*, 1541–1545.

- (11) Stachelska-Wierzchowska, A.; Wierzchowski, J.; Wielgus-Kutrowska, B.; Mikleusević, G. Enzymatic Synthesis of Highly Fluorescent 8-Azapurine Ribosides Using a Purine Nucleoside Phosphorylase Reverse Reaction: Variable Ribosylation Sites. *Molecules* **2013**, *18*, 12587–12598.

- (12) Wierzchowski, J.; Antosiewicz, J. M.; Shugar, D. 8-Azapurines as Isosteric Purine Fluorescent Probes for Nucleic Acid and Enzymatic Research. *Mol. Biosyst.* **2014**, *10*, 2756–2774.

- (13) Da Costa, C. P.; Fedor, M. J.; Scott, L. G. 8-Azaguanine Reporter of Purine Ionization States in Structured RNAs. *J. Am. Chem. Soc.* **2007**, *129*, 3426–3432.

- (14) Stachelska-Wierzchowska, A.; Wierzchowski, J.; Bzowska, A.; Wielgus-Kutrowska, B. Site-Selective Ribosylation of Fluorescent Nucleobase Analogs Using Purine-Nucleoside Phosphorylase as a Catalyst: Effects of Point Mutations. *Molecules* **2016**, *21*, 44–54.

- (15) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discovery* **2015**, *10*, 449–461.

- (16) Wang, C.; Greene, D. A.; Xiao, L.; Qi, R.; Luo, R. Recent Developments and Applications of the MMPBSA Method. *Front. Mol. Biosci.* **2018**, *4*, No. 87.

- (17) Pyrka, M.; Maciejczyk, M. Theoretical Study of Tautomeric Equilibria of 2,6-Diamino-8-Azapurine and 8-Aza-Iso-Guanine. *Chem. Phys. Lett.* **2015**, *627*, 30–35.

- (18) Becke, A. D. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

- (19) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

- (20) Pyrka, M.; Maciejczyk, M. Theoretical Investigations of Tautomeric Equilibrium of 9-Methyl-8-Aza-Iso-Guanine and Its Electrostatic Properties. *Comput. Theor. Chem.* **2016**, *1091*, 1–7.

- (21) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) Theory for Molecules Containing First and Second-Row Atoms. *J. Chem. Phys.* **1998**, *109*, 7764–7776.

- (22) Miertuš, S.; Srocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilization of AB Initio Molecular Potentials for the Prediction of Solvent Effects. *Chem. Phys.* **1981**, *55*, 117–129.

- (23) Miertus, S.; Tomasi, J. Approximate Evaluations of the Electrostatic Free Energy and Internal Energy Changes in Solution Processes. *J. Chem. Phys.* **1982**, *65*, 239–241.
- (24) Toms, A. V.; Wang, W.; Li, Y.; Ganem, B.; Ealick, S. E. Novel Multisubstrate Inhibitors of Mammalian Purine Nucleoside Phosphorylase. *Acta Crystallogr., Sect. D: Struct. Biol.* **2005**, *61*, 1449–1458.
- (25) Bzowska, A.; Koellner, G.; Wielgus-Kutrowska, B.; Stroh, A.; Raszewski, G.; Holý, A.; Steiner, T.; Frank, J. Crystal Structure of Calf Spleen Purine Nucleoside Phosphorylase with Two Full Trimers in the Asymmetric Unit: Important Implications for the Mechanism of Catalysis. *J. Mol. Biol.* **2004**, *342*, 1015–1032.
- (26) Webb, B.; Sali, A. Comparative Protein Modeling Using Modeller. *Curr. Protoc. Bioinf.* **2016**, *54*, 5.6.1–5.6.37.
- (27) Marti-Renom, M. A.; Stuart, A.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. Comparative Protein Structure Modeling of Genes and Genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291–325.
- (28) Sali, A.; Blundell, T. L. Comparative Protein Modeling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
- (29) Fiser, A.; Do, R. K.; Sali, A. Modeling of Loops in Protein Structures. *Protein Sci.* **2000**, *9*, 1753–1773.
- (30) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++3.0: Automating PK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulation. *Nucleic Acids Res.* **2012**, *40*, 537–541.
- (31) Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A. V. A Simple Clustering Algorithm Can Be Accurate Enough for Use in Calculations of PKs in Macromolecules. *Proteins* **2006**, *63*, 928–938.
- (32) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating PKs and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, *33*, W368–W371.
- (33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (34) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An $N \log(N)$ Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (35) Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; Kollman, P. A. *Amber 16*; Technical Report. University of California: San Francisco, 2016.
- (36) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function: Efficient Optimization and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.
- (37) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General AMBER Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (38) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving. *J. Phys. Chem. A* **1993**, *97*, 10269–10280.
- (39) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247–260.
- (40) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A. *LEAP*; University of California: San Francisco, 1995.
- (41) Maier, J. A.; Martinez, C.; Hauser, K.; Kasavajhala, K.; Wickstrom, E.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (42) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.
- (43) Jang, Y. H.; Goddard, W. A., III; Noyes, K. T.; Sowers, L. C.; Hwang, S.; Chung, D. S. PKa Values of Guanine in Water: Density Functional Theory Calculations Combined with Poisson-Boltzmann Continuum-Solvation Model. *J. Phys. Chem. B* **2003**, *107*, 344–357.
- (44) Hanus, M.; Ryjáček, F.; Kabeláč, M.; Kubař, T.; Bogdan, T. V.; Trygubenko, S. A.; Hobza, P. Correlated Ab Initio Study of Nucleic Acid Bases and Their Tautomers in the Gas Phase, in a Microhydrated Environment and in Aqueous Solution. Guanine: Surprising Stabilization of Rare Tautomers in Aqueous Solution. *J. Am. Chem. Soc.* **2003**, *125*, 7678–7688.
- (45) Contreras, J. G.; Madariaga, S. T. Azaguanine: A Theoretical Study of Its Tautomerism and Protonation in the Gas Phase and Aqueous Solution. *Bioorg. Chem.* **1998**, *26*, 345–355.
- (46) Thurlkill, R. L.; Grimsley, G. R.; Scholtz, J. M.; Pace, C. N. PK Values of the Ionizable Groups of Proteins. *Protein Sci.* **2006**, *15*, 1214–1218.
- (47) Yang, T.; Wu, J. C.; Yan, C.; Wang, Y.; Luo, R.; Gonzales, M. B.; Dalby, K. N.; Ren, P. Virtual Screening Using Molecular Simulations. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 1940–1951.
- (48) Wierchowski, J.; Bzowska, A.; Stepniak, K.; Shugar, D. Interactions of Calf Spleen Purine Nucleoside Phosphorylase with 8-Azaguanine, and a Bisubstrate Analogue Inhibitor: Implications for the Reaction Mechanism. *Z. Naturforsch. Sect. C J. Biosci.* **2004**, *59*, 713–725.
- (49) Tan, C.; Tan, Y. H.; Luo, R. Implicit Nonpolar Solvent Models. *J. Phys. Chem. B* **2007**, *111*, 12263–12274.
- (50) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J. Phys. Chem. A* **1994**, *98*, 1978–1988.
- (51) Wang, C.; Nguyen, P. H.; Pham, K.; Huynh, D.; Le, T. N.; Wang, H.; Ren, P.; Luo, R. Calculating Protein – Ligand Binding Affinities with MMPBSA: Method and Error Analysis. *J. Comput. Chem.* **2016**, *37*, 2436–2446.
- (52) Gallicchio, E.; Kubo, M. M.; Levy, R. M. Enthalpy-Entropy and Cavity Decomposition of Alkane Hydration Free Energies: Numerical Results and Implications for Theories of Hydrophobic Solvation. *J. Phys. Chem. B* **2000**, *104*, 6271–6285.