

Dr Zofia Gdaniec

***Wpływ czynników strukturalnych oraz środowiska na
specyficzność nieklasycznych oddziaływań zasada-zasada
w obrębie struktur DNA i RNA. Badania NMR.***

Rozprawa habilitacyjna
przedłożona Radzie Naukowej Instytutu Chemii Bioorganicznej PAN
w Poznaniu

styczeń 2001

BIBLIOTEKA
INSTYTUTU CHEMII BIORGANICZNEJ
POLSKIEJ AKADEMII NAUK
ul. Noskowskiego 12/14
61-704 POZNAŃ

D-213

SPIS TREŚCI

1.	WPROWADZENIE.....	1
2.	STAN RÓWNOWAGI TAUTOMERYCZNEJ N ⁴ -METOKSYCYTOZYNY W DUPLEKSIE DNA ZALEŻY OD TYPU ZASADY PURYNOWEJ W NICI KOMPLEMENTARNEJ ORAZ TEMPERATURY.....	4
3.	NIEKLASYCZNA PARA G-U W OBREBIE WYBRZUSZENIA STRUKTURY 30-meru RNA O SEKWENCJI IRE mRNA DLA FERYTYNY JEST CENTRUM KOMPLEKSOWANIA JONÓW MAGNEZU.....	11
4.	WPLYW N ⁶ -(N-TREONYLOKARBONYLO)ADENOZYNY (t ⁶ A) NA ODDZIAŁYWANIA ZASADA-ZASADA W STRUKTURZE 17-MERU RNA, ANALOGU RAMIENIA ANTYKODONOWEGO tRNA ^{Lys3}	19
5.	UWAGI KOŃCOWE.....	26
6.	CYTOWANA LITERATURA.....	27

1	WPROWADZENIE	1
2	STAN RÓWNOWAGI TAJTOMERYCZNEJ N-METORSYCYTOZYN W DUPLIKSIE DNA ZALEŻY OD TYPU ZASADY PURYNOWEJ W NUCI KOMPLEMENTARNYCH ORAZ TEMPERATURY	4
3	NIEKLASYCZNA PARA G-U W OBRĘBIE WYPRZUSZENIA STRUKTURY 30-mers RNA O SZKWAŃCACH IBE mRNA DŁA NERYTYNY JEST CENTRUM KOMPLEKSOWANIA IONÓW MAGNEZU	11
4	WPLYW N ⁶ -TRETORYLOKARBONYL OXADENOSYNY (T ⁶) NA ODZIAŁYWANIA ZASADY ZASADY W STRUKTURZE I-MEKU RNA ANALOGU RAMIBINA ANTYKOŃCZOWEGO RNA ^{19s}	19
5	UWAGI KOŃCOWE	26
6	CYTOWANA LITERATURA	27

1. WPROWADZENIE

Biomolekularna spektroskopia NMR jest jedyną metodą eksperymentalną pozwalającą na ustalenie współrzędnych atomowych makrocząsteczek w roztworze. Ma to szczególnie duże znaczenie dla rozwiązywania realnych zagadnień biochemicznych np. prób powiązania struktury cząsteczek z ich funkcją w komórce. Zaletą ustalania struktur DNA i RNA metodami NMR jest możliwość rejestrowania funkcjonalnie istotnych zmian konformacyjnych cząsteczki, opis ich dynamiki oraz analiza procesów wiązania ligandów.

Rozwój w dziedzinie spektroskopii NMR kwasów nukleinowych, który dokonał się w ostatnim dziesięcioleciu, zawdzięczamy nie tylko konstrukcji coraz to nowocześniejszych spektrometrów wyposażonych w magnesy o częstotliwościach dla protonów sięgających 900 MHz. Nieustannie pojawiają się nowe techniki pomagające pokonać niedoskonałości metod stosowanych dotychczas w analizie strukturalnej biomolekuł (1-4). Struktury NMR kwasów nukleinowych są otrzymywane z coraz większą precyzją i dokładnością. W tej mierze ustalenia strukturalne dotyczące kwasów nukleinowych i białek metodami biomolekularnego NMR mogą konkurować z najnowszymi osiągnięciami metod krystalograficznych. Ogromny postęp w spektroskopii NMR był również możliwy dzięki opracowaniu metod enzymatycznej syntezy DNA i RNA (5-7) pozwalających na otrzymywanie cząsteczek całkowicie znakowanych izotopami ^{13}C i ^{15}N , oraz postępowi w dziedzinie syntezy chemicznej sekwencyjnie specyficznego znakowania fragmentów RNA (8).

Mam nadzieję, że uzyskane przeze mnie wyniki dyskutowane w komentarzu do rozprawy habilitacyjnej zatytułowanej "*Wpływ czynników strukturalnych oraz środowiska na specyficzność nieklasycznych oddziaływań zasada-zasada w obrębie struktur DNA i RNA. Badania NMR*" mają swój udział w zwiększeniu wiedzy w tej mierze.

Komentarz dotyczy pięciu prac eksperymentalnych:

- I. G.V.Fazakerly, Z.Gdaniec, L.C.Sowers
Base-pair Induced Shifts in the Tautomeric Equilibrium of a Modified DNA Base.
Journal of Molecular Biology, 230, 6-10, 1993

Biolokalizacja spektroskopii NMR jest jedną z metod eksperymentalnych pozwalającą na ustalenie współrzędnych atomowych makrocząstek w roztworze. Ma to szczególne znaczenie dla rozwiązywania problemów z zakresu biochemicznych np. przy poznaniu struktury cząsteczek z ich funkcją w komórce. Zatem ustalenie struktury DNA i RNA metodami NMR jest niezwykle istotnym krokiem w badaniach nad funkcjonowaniem i właściwościami tych związków.

Ważnym w dziedzinie spektroskopii NMR kwasów nukleinowych, który dokonał się w ostatnim dziesięcioleciu, z pewnością jest odkrycie, że tylko konformacja oraz powstawanie i zanikanie punktów wiązania w magnezie o częstotliwościach dla protonów szczytowych 900 MHz. Niezależnie od tego, że nowe techniki pomagające pokonać niedostateczność metod stosowanych dotychczas w analizie strukturalnej biomolekuł (1-4). Szeregi NMR kwasów nukleinowych są niezwykle cenne z punktu widzenia praktyki i doświadczeń. W tej dziedzinie ustalono strukturalnie dotychczas kwasów nukleinowych i białek, myślenie biofizyczne NMR mogą kontynuować z najnowszymi osiągnięciami metod fizykochemicznych. Ograniczony postęp w spektroskopii NMR był również możliwy dzięki opracowaniu metod enzymatycznej syntezy DNA i RNA (5-7) pozwalających na otrzymywanie cząsteczek całkowite znakowanych izotopami ^{13}C i ^{15}N oraz postępowi w dziedzinie syntetycznej chemicznej szwaminy specyficznego trawienia fragmentów RNA (8).

Ważnym zadaniem jest również przez mniejsze wyniki dyskusyjne w komentarzu do rozprawy habilitacyjnej zaplanowanej "Wpływ czynników strukturalnych oraz środowiska na specyficzną selektywność oddziaływań zaindukcji w obszarze struktury DNA i RNA. Badania NMR" mające służyć w dziedzinie wiedzy w tej dziedzinie.

Komentarz dotyczy pięciu prac eksperymentalnych:

1. G.V. Paskov, Z. Góral, L.C. Sorenson
Base-pair induced shift in the Tautomeric Equilibrium of a Modified DNA Base
Journal of Molecular Biology, 120, 5-10, 1989

- II. Z.Gdaniec, B.Baan, L.C.Sowers, G.V.Fazakerley
Methoxyamine-induced Mutagenesis of Nucleic Acids. A Proton NMR Study of Oligonucleotides Containing N⁴-methoxycytosine Paired with Adenine or Guanine.
European Journal of Biochemistry, 242, 271-279, 1996
- III. Z.Gdaniec, H.Sierzputowska-Gracz, E.C.Theil
Iron Regulatory Element (IRE) and Internal Loop/Bulge Structure for Ferritin mRNA Studied by Co(III)hexammine Binding, Molecular Modeling and NMR Spectroscopy.
Biochemistry, 37, 1505-1512, 1998
- IV. Y.Ke, H.Sierzputowska-Gracz, Z.Gdaniec, E.C.Theil
Internal Loop/Bulge and Hairpin Loop of the Iron-Responsive Element of Ferritin mRNA Contribute to Maximal Iron Regulatory Protein 2 Binding and Translational Regulation in the Iso-iron-responsive Element/Iso-iron Regulatory Protein Family.
Biochemistry, 39, 6235-6242, 2000
- V. J.W. Stuart, Z. Gdaniec, R.Guenther, M.Marszałek, E.Sochacka, A.Małkiewicz, P.F.Agris
Functional Anticodon Architecture of Human tRNA^{Lys3} Includes Disruption of Intraloop Hydrogen Bonding by the Naturally Occurring Amino Acid Modification, t⁶A.
Biochemistry, 39, 13396-13404, 2000

W komentarzu koncentruję się na dyskusji aspektów strukturalnych nieklasycznych oddziaływań zasada-zasada zarówno w kontekście tworzenia wiązań wodorowych jak i oddziaływań warstwowych. Wskazany powyżej stan metod syntezy chemicznej i enzymatycznej fragmentów DNA i RNA oraz rozwój spektroskopii NMR umożliwił mi wniknięcie w naturę nieklasycznych oddziaływań na poziomie atomowym w roztworze.

Aby uniknąć nieścisłości dotyczących nomenklatury chcę dodać, że wyrażeniem „nieklasyczne oddziaływania” objęłam te, które prowadzą do powstania par zasad innych niż kanoniczne, postulowanych przez Watsona i Cricka. W moim komentarzu unikam stwierdzeń typu „niedopasowane” czy „niesparowane” zasady, jako odpowiednika angielskich wyrażen *mismatch* i *mispair*. Jak bowiem wskażę, tzw. niedopasowane zasady są często lokowane (dopasowywane?) w obrębie struktur DNA czy RNA w taki sposób, aby indukowane nimi zaburzenie było jak najmniejsze.

- II. Z. Górniewicz, H. Banaś, L. C. Sawyer, D. V. Frenkel
Methicillin-induced Alignment of Nucleic Acids: A Proton NMR Study of
Oligonucleotides Containing N-methylcytosines Paired with Adenine or Guanine.
European Journal of Biochemistry, 242, 271-279, 1997
- III. Z. Górniewicz, H. Sierżputowska-Górniewicz, E. C. Tjell
Iron Regulatory Element (IRE) and Internal Loop/Bulge Structure for Ferritin
mRNA Studied by Co(III)-mediated Binding, Molecular Modeling and NMR
Spectroscopy.
Biochemistry, 37, 1505-1512, 1998
- IV. Y. Ke, H. Sierżputowska-Górniewicz, Z. Górniewicz, E. C. Tjell
Internal Loop/Bulge and Hoop Loop of the Iron-Responsive Element of Ferritin
mRNA Contribute to Minimal Iron Regulatory Protein 2 Binding and Translational
Regulation in the Iron-Responsive Element-to-Iron Regulatory Protein Family.
Biochemistry, 39, 6233-6242, 2000
- V. J. W. Sauer, Z. Górniewicz, R. Günther, M. Matuszewska, E. Sobczak, A. Malinowska,
T. A. Agre
Functional Backbone Architecture of Human tRNA^{Met} Involves Distortion of
Internal Loop/Bulge Binding by the Nucleoside Occurring Amino Acid
Modification Y4.
Biochemistry, 39, 13396-13404, 2000

W komentarzu koncentruje się na dyskusji aspektów strukturalnych niekwalifikacji oddziaływań zasady-zasada zarówno w kontekście tworzenia wiązań wodorowych jak i oddziaływań wariantowych. Wskazany powyżej stan metod syntezy chemicznej i enzymatycznej fragmentów DNA i RNA oraz rozwój spektroskopii NMR umożliwił mi właśnie w naturę niekwalifikacji oddziaływań na poziomie atomowym w rozwarze. Aby uzyskać niekwalifikacji dotyczących komplementarych chęć dodać, że wyrażeniem „niekwalifikacja oddziaływań” opisana jest, które prowadzi do powstania par zasad innych niż kanoniczne, postulowanych przez Watsona i Cricka. W moim komentarzu ukażą się również typy „niekwalifikowane”, czy „niekwalifikowane”, zasady, jako odpowiednia angielskich wyrazów *mispaired* i *mismatch*. Jak powinien wyglądać tzw. niekwalifikowane zasady są często lokowane (dopasowywane) w odczynie struktur DNA czy RNA w taki sposób, aby indukowane nimi zaburzenie było jak najmniej.

Moje obserwacje poczynione metodami biomolekularnego NMR dotyczą zarówno nieklasycznych oddziaływań obejmujących kanoniczne zasady np. G-U w cząsteczkach RNA, jak i tych obejmujących modyfikowane nukleozydy w tRNA, czy też produkty mutagenyzy chemicznej DNA.

Różnorodność form strukturalnych obserwowana dla funkcjonalnie ważnych domen DNA i RNA w roztworze jest wypadkową wpływu oddziaływań wewnątrzcząsteczkowych oraz warunków środowiska. We wszystkich przypadkach obserwowałam ustalanie się swoistego stanu równowagi, który był rezultatem wpływów zaburzeń, jakie nieklasyczne pary zasad wywierają na globalną strukturę cząsteczki DNA czy RNA, oraz odpowiedzią całej cząsteczki na obecność centrum modyfikacji.

Moje opracowanie poświęcone metodami biologicznym NMR dotyczy zarówno nielazycznych oddziaływań obejmujących kanoniczne zasady np. G-U w cząsteczkach RNA, jak i tych obejmujących modyfikowane nukleotydy w tRNA, czy też produkty mutageny chemicznej DNA.

Różnorodność form strukturalnych obserwowana dla funkcjonalnie ważnych dimerów DNA i RNA w roztworze jest wypadkową wpływu oddziaływań wewnątrzcząsteczkowych oraz warunków środowiska. We wszystkich przypadkach opracowania ustalają nie tylko stan równowagi, który był rezultatem wpływów zewnętrznych, jak również wywierają na globalną strukturę cząsteczki DNA czy RNA, oraz odpowiednio całej cząsteczki na obecność centrum modyfikacji.

2. STAN RÓWNOWAGI TAUTOMERYCZNEJ N⁴-METOKSYCYTOZINY W DUPEKSIIE DNA ZALEŻY OD TYPU ZASADY PURYNOWEJ W NICI KOMPLEMENTARNEJ ORAZ TEMPERATURY.

Niekanoniczne pary zasad mogą powstawać *in vivo* w strukturze dwuniciowego DNA najczęściej w wyniku błędów w procesie replikacji lub w trakcie rekombinacji pomiędzy homologicznymi, ale nieidentycznymi sekwencjami DNA. W procesie replikacji nieuchronnie pojawiają się błędy, najczęściej mutacje punktowe: tranzycja i transwersja. Innym typem mutacji jest delecja lub insercja jednej lub więcej par zasad.

Oprócz błędów, w których uczestniczą tylko zasady kanoniczne, mutacje mogą być wywołane także przez chemiczne modyfikacje zasad DNA. Doprowadza to w konsekwencji do zmiany struktury elektronowej nukleozasady, a tym samym w sposób zasadniczy zmienia schemat kanonicznego formowania par A-T i C-G. Pośród wielu mutagenów chemicznych częstym i istotnym czynnikiem prowadzącym do modyfikacji zasad jest atak rodników hydroksylowych (9), tworzących się w roztworach wodnych pod wpływem promieniowania jonizującego.

Powstające błędy mogą zostać naprawione dzięki niezwyklej, dodatkowej aktywności polimeraz DNA, która polega na sprawdzaniu poprawności wbudowywania nukleotydów podczas replikacji (10). Błędy, które wymknęły się pierwszemu procesowi naprawy przez polimerazę DNA, poddawane są następnie działaniu systemów naprawczych *in vivo* (10-12). Ta podwójna linia obrony sprawia, że po skończonej akcji wszystkich systemów naprawczych pozostaje tylko jeden błąd na 10^{11} - 10^{12} par zasad w nowo zsyntetyzowanej nici DNA. Postreplikacyjny, enzymatyczny system naprawczy rozpoznaje i naprawia nie tylko nieklasyczne pary zasad. Naprawie podlegają również chemicznie modyfikowane pary zasad. Niepowodzenie podczas ich naprawy prowadzi do przekazania komórkom potomnym mutacji potencjalnie skutkujących nowotworzeniem lub chorobą genetyczną. Z tych względów zrozumienie na poziomie strukturalnym mechanizmu rozpoznania i naprawy struktur DNA zawierających niekanoniczne pary zasad wymaga nadal szczegółowych badań.

Datowane sprzed okresu 1985-1990 badania za pomocą spektroskopii NMR dotyczące właściwości chemicznie modyfikowanych zasad w obrębie struktur DNA, były niezwykle utrudnione. Źródłem tego były ograniczenia w syntezie chemicznej i

Niekonwencjonalne pary zasad mogą powstawać in vivo w strukturze dwuniciowego DNA najczęściej w wyniku błędów w procesie replikacji lub w trakcie rekombinacji genetycznej, homologicznej, ale niebędącej reasortacją DNA. W procesie replikacji nieuchronnie pojawiają się błędy, najczęściej mutacje punktowe: transycja i transwersja. Innym typem mutacji jest delecja lub insercja jednego lub więcej par zasad.

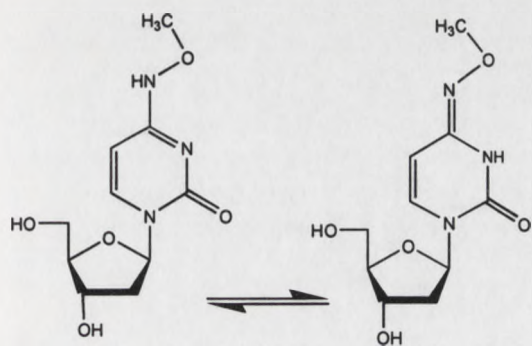
Opisując błędy, w których uczestniczą tylko zasady kanoniczne, mutacje mogą być wywołane także przez obniżenie energii wiązania zasad DNA. Dotyczy to w szczególności do zmiany struktury siatkowej nukleozymu, a tym samym w sposób zasadniczy zmienia schemat kanonicznego formowania par A-T i C-G. Pośród wielu mutacji chemicznych, takich jak i inne, w tym mutacje indukowane, do mutacji należą zasadnie takie rodzaje błędów (B), które powstają w wodorowych wiązaniach między komplementarnymi zasadami.

Powstałe błędy mogą zostać naprawione dzięki niezwykłej zdolności aktywności polimerazy DNA, która polega na sprawdzaniu poprawności wodorowych wiązanych podczas replikacji (10). Błędy, które wynikły z nieprawidłowego naprawy przez polimerazę DNA, poddawane są następnemu działaniu systemów naprawczych in vivo (10-12). Ta podwójna linia obrony sprawia, że po skończonej akcji wszystkich systemów naprawczych pozostaje tylko jeden błąd na 10^{11} - 10^{12} par zasad w nowo syntetyzowanej nici DNA. Ponieważ błędy, szczególnie te, które powstają w toku replikacji i naprawy nie tylko niekiedy par zasad. Nierzadko dochodzi do chemicznej modyfikacji par zasad. Niepewność podczas ich naprawy prowadzi do przekazania kodu błędów potomnym komórkom. Wzrost liczby błędów skutkuje nowotworzeniem lub chorobą genetyczną. Z tych względów zwracanie na uwagę struktur DNA przy mechanizmie rozpoznania i naprawy struktury DNA zawierających niekonwencjonalne pary zasad wymaga nadal szczególnych badań.

Długość czasu okresu 1982-1990 badania za pomocą spektroskopii NMR dotyczące właściwości chemicznej modyfikowanych zasad w obrębie struktury DNA były niezwykle istotne. Zdobytym takim były ograniczenia w syntezie chemicznej i

oczyszczaniu miligramowych ilości modyfikowanych oligodeoksynukleotydów, oraz w metodyce NMR. Zdecydowana większość modyfikacji przyczynia się do destabilizacji cząsteczki DNA. Obecność mutacji punktowej o rozbudowanej strukturze może doprowadzić do obniżenia tzw. temperatury topnienia (T_m) oligonukleotydu nawet o 20°C (13). Obecność dwóch lub więcej destabilizujących modyfikacji w krótkich wzajemnie komplementarnych oligodeoksynukleotydach prowadzi zazwyczaj do układów o małej stabilności termicznej, co uniemożliwia interpretację wpływu modyfikacji metodami spektroskopii NMR (14). Z tych względów w swoich badaniach stosowałam jedynie duplekisy DNA tworzące się w wyniku asocjacji dwóch nici, z których żadna nie posiada właściwości wzajemnej komplementacji. Badania ilościowe za pomocą metod NMR mogą być wykonane tylko w przypadku, gdy obserwowane w widmach linie rezonansowe nie są wynikiem procesu wymiany pomiędzy strukturą jedno- i dwuniciową. Dla krótkich oligomerów wykazujących tendencję do tworzenia niekanonicznych par, dobranie warunków pomiarowych tak, aby istniała tylko struktura dupleksu, może okazać się niemożliwe wobec potencjalnie niekorzystnego wpływu czynników takich jak pH, temperatura, sekwencja zasad czy też obecność różnych jonów w roztworze. Niezwykle istotne jest więc poznanie wpływu tych parametrów na badany układ, w celu dobrania optymalnych warunków eksperymentalnych.

Hydroksyloamina oraz jej analog metoksyamina, są znanymi mutagenami, które *in vivo* generują tranzycje z C-G do T-A (15). Wynika to z odmiennego parowania zasad w kolejnym cyklu replikacji, następującym po zmodyfikowaniu zasady. W efekcie działania metoksyaminy na deoksycytydynę lub eksponowane reszty cytozynowe DNA dochodzi do zastąpienia jednego z atomów wodoru grupy egzoaminowej grupą metoksyłową. W odróżnieniu od deoksycytydyny, której zdecydowanie uprzywilejowaną formą tautomeryczną jest forma aminowa, N⁴-metoksytydyna (mo⁴C) występuje w roztworze

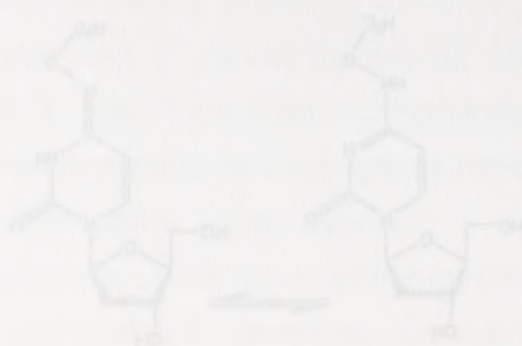


głównie jako tautomer iminowy z grupą metoksyłową zorientowaną *syn* w stosunku do atomu azotu N-3 zasady (16-18) (patrz rysunek). Różnica energetyczna pomiędzy formą iminową a aminową dla monomeru mo⁴C jest jednak bardzo niska (17,19).

oczyszczaniu modyfikowanych linii modyfikowanych oligonukleotydów, oraz w metodach NMR. Zbadano wpływ wielkości modyfikacji przytoczonej się do destabilizacji cząsteczki DNA. Obecność mutacji punktowej o rozbudowanej strukturze może doprowadzić do obniżenia tzw. temperatury topnienia (T_m) oligonukleotydu nawet o 20°C (13). Obecność dwóch lub więcej destabilizujących modyfikacji w krótkich wzajemnie komplementarnych oligonukleotydach prowadzi zazwyczaj do układów o niskiej stabilności termicznej, co umożliwia interakcję wpływu modyfikacji metodami spektroskopii NMR (14). Z tych względów w swoich badaniach stosowałam jedynie dupleksy DNA tworzące się w wyniku kopacji dwóch nici, z których każda nie posiada własności wzajemnej komplementarności. Badanie ilościowe za pomocą metod NMR może być wykonane tylko w przypadku, gdy obserwowane w widmach linie rezonansowe nie są wynikiem procesu wymiany pomiędzy strukturami jedno- i dwuniciowymi. Dla krótkich oligonukleotydów tendencją do tworzenia niekanonicznych par, dobrane warunków pomiarowych tak, aby istniała tylko jedna struktura dupleksu, może okazać się niemożliwe wobec potężnego niekanonicznego wpływu czynników takich jak pH, temperatura, składowe zasady czy też obecność jonów w roztworze. Należy pamiętać, że istotne jest również poznanie wpływu tych parametrów na badany układ, w celu dobrania optymalnych warunków eksperymentalnych.

Hydroksylamina oraz jej analogi modyfikują, są mianem mutagenami, które *in vivo* generują mutacje z C-G do T-A (15). Wynika to z odmianowego powstania zasad w kolejnym cyklu replikacji, następującego po zmodyfikowaniu zasady. W etapie działania metylaminu na deoksyrybozę lub odpowiednio na cytozynowe DNA dochodzi do zastąpienia jednego z atomów wodoru grupy cytozynowej grupą metylową. W oddziaływaniu od deoksyrybozy, której zbadanie ukazywało formę tautomerową jest formą aminową. N-deksozyryboza (m^+) występuje w roztworze głównie jako tautomer iminowy z grupą

aminoową zorientowaną *syn* w stosunku do atomu azotu N-3 zasady (16-18) (patrz rysunek). Różnica energetyczna pomiędzy formą iminową a aminową dla monomeru m^+ jest jednak bardzo niska (17,19).



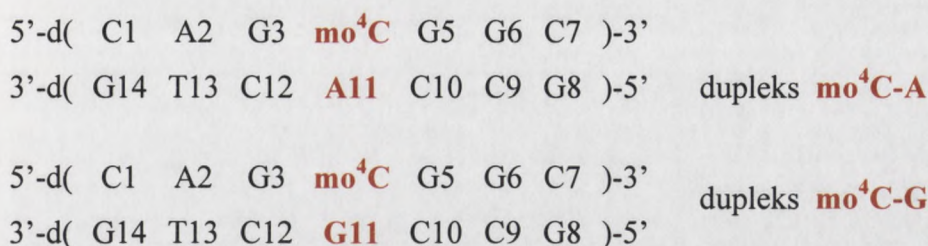
Badania strukturalne nukleotydów, których celem było wyjaśnienie mechanizmu mutacji wywołanych działaniem metoksyaminy, zapoczątkowano w grupie Shugara na początku lat 80-tych (20,21). Dużą częstość występowania tranzycji z C-G do T-A przypisywano od samego początku niejednoznaczny właściwościom m^4C w aspekcie parowania z A i G. Wykonane znacznie później pomiary stabilności dupleksów zawierających m^4C (14,22) wykazały, że N^4 -metoksytydyna może tworzyć stabilne pary zasad zarówno z deoksyadenozyną jak i deoksyguanozyną.

W początkowej fazie moich badań ukazała się praca z laboratorium Brown'a (23), w której do badań NMR użyto oligodeoksynukleotydy o charakterze wzajemnie komplementarnym, zawierające modyfikację m^4C , a więc obarczone powyżej opisaną ułomnością. Badania te dla wzajemnie komplementarnego oktameru nie dały jednoznacznych wyników. Obecność dwóch modyfikowanych par zasad w strukturze duplesu przyczyniła się do jego destabilizacji. Temperatura topnienia obniżyła się o około $20^\circ C$ w stosunku do analogicznego duplesu zawierającego kanoniczne pary zasad A-T i wynosiła $33^\circ C$. Złożoność widm NMR zarejestrowanych dla duplesu posiadającego pary m^4C -A wynikała zarówno z poszerzenia wielu sygnałów rezonansowych jak i pojawienia się dodatkowych. Autorzy tłumaczyli to istnieniem skomplikowanej równowagi pomiędzy wieloma formami. Postulowali oni, mimo braku bezpośrednich dowodów eksperymentalnych, wolny w skali czasu NMR proces wymiany pomiędzy strukturami zawierającymi pary zasad m^4C -A w geometrii Watsona-Cricka i kilkoma typami par *wobble*. Oprócz różnych form tautomerycznych m^4C autorzy sugerowali dodatkowo, że dla par typu *wobble* w równowadze znajdowały się także geometryczne izomery (*syn-anti*) reszt m^4C . Podobne wyniki otrzymano dla analogicznego duplesu, ale z parami m^4C -G (24). Również i w tym przypadku obserwowano wiele różnych form pozostających ze sobą w równowadze.

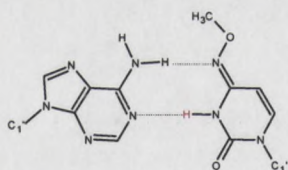
Tendencja N^4 -metoksytyozyny do tworzenia par z adeniną i guaniną stanowiła dla mnie niezwykle okazję do eksperymentalnego udowodnienia zależności wpływu czynników wewnątrzstrukturalnych na stan równowagi amino-iminowej m^4C oraz na tworzenie się w roztworze określonego typu pary zasad w dupleksach DNA. Prowadzone przeze mnie badania NMR w laboratorium dr G.V. Fazakerley'ego (CEA, CEN Saclay, Francja) obejmowały dwa dupleksy DNA o zasadniczo tej samej sekwencji. W centralnej pozycji naprzeciwko pojedynczej reszty m^4C umieszczono adeninę lub guaninę (I).

Balans...
mutacji...
postaci...
przy...
prowiz...
zawiesz...
przy...
W...
w...
komplement...
niom...
jednost...
duplik...
20°C...
wyno...
m...
się...
w...
okre...
zawiesz...
w...
dla...
m...
(24) R...
w...

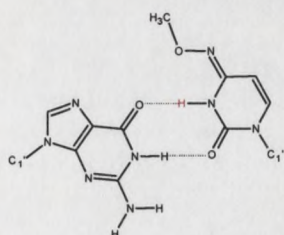
Tendencja...
nie...
cy...
twor...
prze...
Fr...
pozy...



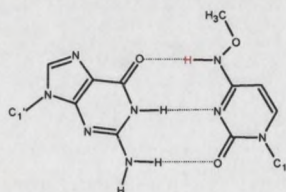
W przeciwieństwie do omówionych powyżej wyników, rezultaty otrzymane przeze mnie dla dupleksu zawierającego nieklasyczną parę mo^4C -A (**I**), pokazały istnienie w roztworze tylko jednej formy tautomerycznej mo^4C . Na podstawie analizy widm NOESY stwierdziłam, że badany dupleks posiadał regularną formę B-DNA. Centralna para zasad miała geometrię typu Watsona-Cricka, w której reszta mo^4C występowała jako tautomer iminowy. Wymogiem strukturalnym dla utworzenia takiej pary pomiędzy mo^4C a resztą adeniny jest nie tylko obecność iminowej formy tautomerycznej mo^4C , ale również orientacja *anti* grupy metoksylowej. Zmiana preferowanej, obserwowanej na poziomie monomeru oraz pojedynczej nici, orientacji grupy metoksylowej z *syn* do *anti* w parze mo^4C -A jest czynnikiem spowalniającym proces wymiany pomiędzy strukturą dupleksu a formą jednoniciową. Proces ten jest wolny w skali czasu NMR, co zaobserwowałam rejestrując przebieg krzywych topnienia dla dupleksu mo^4C -A (**II**). W temperaturze 37°C pojawiały się nowe sygnały od struktury jednoniciowej. W temperaturze 47°C populacja obu form, dupleksu i pojedynczej nici, była równa.



para A- mo^4C w geometrii Watsona-Cricka,
grupa metoksylowa *anti*



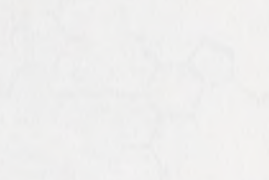
para G- mo^4C w geometrii *wobble*
grupa metoksylowa *syn*



para G- mo^4C w geometrii Watsona-Cricka,
grupa metoksylowa *anti*

2-41 CI 101 111 121 131 141 151 161 171 181 191 201
 2-42 CI 101 111 121 131 141 151 161 171 181 191 201
 2-43 CI 101 111 121 131 141 151 161 171 181 191 201
 2-44 CI 101 111 121 131 141 151 161 171 181 191 201

The following table shows the results of the experiments conducted on the various samples. The data indicates that the reaction rate is significantly higher for the samples with a higher surface area, as expected. The results also show that the reaction rate is affected by the concentration of the reactants, with higher concentrations leading to a faster reaction rate. The data further demonstrates that the reaction rate is also influenced by the temperature, with higher temperatures resulting in a faster reaction rate. The overall findings suggest that the reaction rate is a function of the surface area, concentration, and temperature of the reactants.



The diagram shows the reaction mechanism for the formation of the product. The reactants are shown on the left, and the product is shown on the right. The reaction proceeds through a series of steps, with the transition state shown in the middle. The diagram highlights the key steps in the reaction, including the formation of the intermediate and the final product.

Badania NMR dupleksu z nieklasyczną parą zasad m^4C-G dały zaskakujące rezultaty (I,II). W roztworze zaobserwowałam istnienie dwóch różnych form, będących w równowadze. Na podstawie analizy widm NMR ustaliłam, że w tworzeniu tych dwuniciowych struktur udział biorą oba tautomery reszty m^4C . Gdy m^4C jest w formie aminowej, to powstająca para zasad m^4C-G ma geometrię typu Watsona-Cricka, natomiast forma iminowa N^4 -metoksycytozyny wymusza strukturę parowania zasad typu *wobble*. Ustalająca się równowaga pomiędzy tymi formami zależy od temperatury. W wyższych temperaturach dominującą formą m^4C jest forma iminowa, natomiast w niższych temperaturach przeważa jej tautomer aminowy. Sygnały pochodzące od aminowej formy dupleksu m^4C-G były widoczne w widmie do temperatury około $40^{\circ}C$.

Wykorzystując metody spektroskopii NMR przeprowadziłam badania kinetyczne tych przejść (II). Badań tych nie mogłam prowadzić w dużym zakresie temperatur, gdyż z jednej strony przeszkodą był zbyt wolny proces wymiany w niższych temperaturach, z drugiej strony ograniczona byłam temperaturą topnienia dupleksu. Otrzymane w temperaturze $30^{\circ}C$ dane wskazują na model dwustanowy ze stałą szybkości wymiany pomiędzy formami wynoszącą około jednej sekundy.

W tym miejscu warto podkreślić, że w sytuacji, gdy krótkie oligonukleotydy zawierają tylko pary zasad typu Watsona-Cricka, wymiana pomiędzy strukturą pojedynczej nici i podwójną helisą jest szybka w skali czasu NMR. Proces wymiany ulega spowolnieniu w przypadku dłuższych oligonukleotydów. Dla pewnych układów z niekanonicznymi lub posiadającymi modyfikacje parami zasad, proces wymiany może być wolniejszy nawet o kilka rzędów wielkości. Na ogół szybkość tej wymiany jest zdeterminowana poprzez uprzywilejowaną konformację pojedynczej nici. Należy pamiętać, że nawet w przypadku krótkich, jednoniciowych oligonukleotydów możemy mieć do czynienia ze znacznym ustrukturalizowaniem oligomeru, który na ogół przyjmuje postać prawoskrętnej helisy.

Analizując krzywe topnienia dupleksu m^4C-G byłam w stanie opisać tworzenie formy iminowej i aminowej (II). Dla formy iminowej obserwowałam typowe krzywe sigmoidalne. W odróżnieniu od m^4C-A , tworzenie się dupleksu z iminową formą pary m^4C-G nie wymaga zmiany orientacji grupy metoksyłowej z *syn* do *anti* i dlatego w tym przypadku wymiana pomiędzy formą jednoniciową a dupleksem jest szybka w skali czasu NMR.

Badania NMR dwulekton z niekierującą parą zasad $\text{m}^2\text{-C-G}$ daly zaskakujące rezultaty (II). W roztworze zaobserwowaliśmy istnienie dwóch różnych form, będących w równowadze. Na podstawie analizy widm NMR ustaliliśmy, że w tworzonym tych dwulektonowych struktur udział bierze para zasad $\text{m}^2\text{-C-G}$. Gdy $\text{m}^2\text{-C}$ jest w formie aminowej, to powstaje para zasad $\text{m}^2\text{-C-G}$ ma geometrię typu Watson-Cricka, natomiast forma iminowa N^2 -metylocytozyny wymaga struktur parowania zasad typu wosła. Ustalenie się równowagi pomiędzy tymi formami zależy od temperatury. W wyższych temperaturach dominuje forma $\text{m}^2\text{-C}$ jest formą iminową, natomiast w niższych temperaturach przeważa jej tetramer aminowy. Sygnały pochodzące od aminowej formy dwulekton $\text{m}^2\text{-C-G}$ były widoczne w widmie do temperatury około 40°C .

Wykorzystując metody spektroskopii NMR przeprowadziliśmy badania kinetyczne tej pary (II). Badania te nie mogły prowadzić w dużym zakresie temperatur, gdyż z jednej strony przeszkadzał był zbyt wolny proces wymiany w niższych temperaturach, z drugiej strony ograniczona była temperatura topnienia dwulektona. Otrzymał w temperaturze 30°C dane wskazujące na model dwulektonowy, ze stałą szybkości wymiany pomiędzy formami wynoszącej około jednej sekundy.

W tym miejscu warto podkreślić, że w sytuacji gdy źródło oligonukleotydów zawiera tylko parę zasad typu Watson-Cricka, wymiana pomiędzy strukturami pojedynczej nici i parującym dwulektonem jest szybka w skali czasu NMR. Proces wymiany polega spowolnieniu w przypadku danych oligonukleotydów. Dla parowej układowi z niekierującą parą lub posiadającą inne metody (np. parami zasad, proces wymiany może być wolniejszy nawet o kilka rzędów wielkości. Na ogół szybkość tej wymiany jest zdecydowanie powolniejsza niż w przypadku kontynuacji pojedynczej nici. Należy pamiętać, że nawet w przypadku krótkich, jednoniciowych oligonukleotydów możemy mieć do dyspozycji ze znacznym asynchronizowaniem oligomera, który na ogół przyjmując postać przewoźniczej helisy.

Analizując krzywe topnienia dwulekton $\text{m}^2\text{-C-G}$ byłam w stanie opisać tworzenie formy iminowej i aminowej (II). Dla formy iminowej obserwowaliśmy typowe krzywe sigmoidalne. W odróżnieniu od $\text{m}^2\text{-C-A}$, tworzenie się dwulekton z trójną formą parę $\text{m}^2\text{-C-G}$ nie wymaga zmiany orientacji grupy metaksoylowej aż do zwrotu i dlatego w tym przypadku wymiana pomiędzy formami jednoniciową a dwulektonem jest szybka w skali czasu

Przeciwnie, dla zawiązania aminowej formy pary mo^4C-G nieodzowne są aż dwie zmiany w strukturze mo^4C - reorientacja grupy metoksylovej z *syn* do *anti* oraz tautomeryzacja z formy imino do amino. Zmiany te przyczyniają się do spowolnienia procesu wymiany chemicznej pomiędzy strukturą jedno- a dwuniciową. Zjawisko to zaobserwowałam jako poszerzenie się sygnałów w widmach NMR w temperaturach bliskich 40°C.

Wyniki badań w roztworze często odnosimy do tych otrzymanych w kryształach. Te ostatnie zazwyczaj traktuje się jako nadrzędne, dlatego chciałabym ustosunkować się do rezultatów otrzymanych w kryształach dla badanych przeze mnie modyfikacji. W 1990 roku opublikowano (18) strukturę wzajemnie komplementarnego dupleksu $d(CGCGmo^4CG)_2$. W kryształach pary mo^4C-G posiadały geometrię typu *wobble*. Reszty mo^4C występowały w formie iminowej, a grupa metoksylovej była zorientowana *syn* w stosunku do atomu azotu N-3 zasady. Należy podkreślić, że dupleks ten wykrył się w formie Z-DNA, trudno więc stwierdzić, jak dalece tworzenie lewoskrętnej struktury dupleksu wpłynęło na przestrzenną modyfikację pary mo^4C-G . Badania kryształu wzajemnie komplementarnego dodekameru w formie B-DNA zawierającego parę mo^4C-A , a nie parę mo^4C-G jak powyżej, zostały opublikowane dopiero w 1999 roku (25). W strukturze tej mo^4C występuje jako tautomer iminowy z grupą metoksylovej zorientowaną *anti* i tworzy z resztą adeniny parę o geometrii Watsona-Cricka. Uważam, że w przypadku interesujących mnie oddziaływań, wybór obiektów do badań w kryształach (wzajemnie komplementarne dupleksy DNA niosące dwie modyfikowane pary zasad), nie był optymalny.

W moich badaniach zdołałam wykazać (I,II), że w roztworze poziom równowagi tautomerycznej pomiędzy formą iminową a aminową N^4 -metoksyocytozyny w dupleksach B-DNA jest zależny zarówno od rodzaju zasady purynowej w nici komplementarnej jak i od temperatury. Mamy więc do czynienia ze swoistym zjawiskiem indukcji takiej formy mo^4C , która w sposób najpełniejszy umożliwia stabilizację oddziaływania zasada-zasada. W przypadku tworzącej się pary mo^4C-A o geometrii Watsona-Cricka jest to forma aminowa, która wymusza orientację *anti* grupy metoksylovej. W przypadku pary zasad mo^4C-G , w roztworze istnieją w równowadze dwie różne formy, w tworzeniu których udział biorą oba tautomery reszty N^4 -metoksyocytozyny. Gdy mo^4C jest w formie aminowej, to powstająca para zasad mo^4C-G ma geometrię typu Watsona-Cricka z grupą

Przewidywanie, dla zawiązania anionowej formy pary m^+C-G nieobchodzone są aż dwie zmiany w strukturze m^+C - konieczność grupy metylowej z tym do woli oraz izostereizacja z formą imino do anion. Zmiany te przekazywały się do spowodowania procesu wymiany chemicznej pomiędzy strukturą jedno- a dwucieczną. Zjawisko to obserwowaliśmy jako powstawanie się sygnałów w widmach NMR w temperaturach bliskich $40^\circ C$.

Wymibi badań w roztworze często odnosiły do tych otrzymanych w kryształ. Te ostatnie zawięzają kształtę się jako narządki, dlatego chcieliśmy zorientować się do rezultatów otrzymanych w kryształ dla badania przez nas modyfikacji. W 1990 roku opublikowano (18) strukturę wzajemnie komplementarnego dupleksu (m^+C-G) . W kryształ pary m^+C-G posiadały geometryczną formę wolną. Reakty m^+C występowały w formie iminowej, a grupa metylowa była zorientowana przy w kierunku do atomu azotu N-3 zasady. Należy podkreślić, że dupleks ten wytworzył w formie Z-DNA, trudno więc stwierdzić, jak dalece tworzenie kwadrupletu struktury dupleksu wpłynęło na przekształcenie modyfikacji pary m^+C-G . Badania kryształ wzajemnie komplementarnego dupleksu w formie B-DNA trójczłonek pary m^+C-A , a nie pary m^+C-G jak powyżej, zostały opublikowane dopiero w 1999 roku (22). W strukturze tej m^+C występowało jako anion iminowy z grupą metylową zorientowaną tam i tworzy z ciałem adenylną parę o geometrii Watsona-Cricka. Uwolniam, że w przypadku interakcji takie oddziaływały, wybór obiektów do badań w kryształ (wzajemnie komplementarne dupleksy DNA nie sąc dwie modyfikowane pary zasad) nie był optymalny.

W moich badaniach zdołałem wystrząć (11), że w roztworze poziom równowagi autonomicznej pomiędzy formą iminową a anionową N^+ -metylocytozyny w dupleksach B-DNA jest zależny zarówno od rodzaju zasady purynowej w nici komplementarnej jak i od temperatury. Mały wpływ do czynienia ze wzrostem zjawiskiem indukcji tędy formę m^+C krótu w sposób niepożądany zmniejsza stabilizację oddziaływanie zasady-zasada. W przypadku tworzenia się pary m^+C-A o geometrii Watsona-Cricka jest to forma anionowa krótu wymusza orientację tam grupy metylowej. W przypadku pary zasad m^+C-G , w roztworze istnieją w równowadze dwie różne formy, w tworzeniu krótu udział biorą obie autonomicznej formy N^+ -metylocytozyny. Gdy m^+C jest w formie iminowej, to powstaje para zasad m^+C-G o geometrii typu Watsona-Cricka z grupą

metoksyłową zorientowaną *anti*, natomiast forma iminowa N⁴-metoksyctozyny wymusza geometrię parowania zasad typu *wobble*, dla której możliwe jest zachowanie uprzywilejowanej konformacji *syn* grupy metoksyłowej.

Przedstawione wyniki strukturalne wskazujące na to, że mo⁴C może tworzyć stabilne, niekanoniczne pary zasad zarówno z adeniną jak i guaniną, tłumaczą obserwacje genetyczne o tendencji reszt mo⁴C do indukowania tranzycji.

metakrylowy zawierający oraz kationem formy iminowej N⁺-metakryloylowy w wyniku geometrycznego przewrotu zasad typu wodoru, dla której możliwe jest zachowanie przyswajalności kationami z grupy metakrylowej.

Przedstawione wyniki strukturalne wskazują na to, że m⁺C może tworzyć stabilne, niekanonyczne pary zasad zarówno z adeniną jak i guaniną, tłumaczy obserwacje genetyczne o tendencji term m⁺C do indukowania mutacji.

3. NIEKLASYCZNA PARA G-U W OBRĘBIE WYBRZUSZENIA STRUKTURY 30-meru RNA O SEKWENCJI IRE mRNA DLA FERYTYNY JEST CENTRUM KOMPLEKSOWANIA JONÓW MAGNEZU.

Spośród wszystkich klas biopolimerów, cząsteczki RNA najtrudniej poddają się analizie NMR. W widmach ^1H NMR oligorybonukleotydów sygnały protonów H2', H3', H4', H5' i H5'' skupiają się w bardzo wąskim zakresie (ca. 1.5 ppm). Prowadzi to do silnego nakładania się linii rezonansowych, znacznie większego niż w przypadku widm DNA czy białek. Pełna analiza widm NMR fragmentów RNA o masie cząsteczkowej rzędu 10kD (~30 reszt nukleotydowych) jest niemożliwa bez zastosowania znakowania izotopowego $^{13}\text{C}/^{15}\text{N}$. Źródłem szczególnych trudności w badaniach strukturalnych RNA są ich funkcjonalnie istotne regiony jednoniciowe wykazujące znaczną dynamikę strukturalną. Ustalenie struktury takiego fragmentu RNA jest niezwykle trudne, a często wręcz niewykonalne. Dużą niedogodność w analizie strukturalnej fragmentów RNA stanowi również ich szczególna podatność na degradację wywoływaną przez wszechobecne nukleazy czy też hydrolizę promowaną jonami metali.

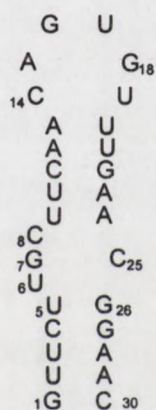
Chociaż opisane w tym rozdziale badania miały początkowo na celu uzupełnienie i poszerzenie informacji zgromadzonych wcześniej (26,27) w laboratorium dr. E. Theil (NCSU, Raleigh, USA) o strukturze fragmentu RNA o sekwencji IRE, to w rezultacie doprowadziły mnie do interesujących ustaleń dotyczących nieklasycznych oddziaływań w obrębie tej cząsteczki. Badana przeze mnie cząsteczka IRE o długości 30 reszt nukleotydowych posiadała sekwencję naturalnie występującą w mRNA dla podjednostki H ferytyny z *Rana catesbeiana* (gatunek żaby północnoamerykańskiej). Elementy strukturalne typu IRE (ang. *Iron Responsive Element*) są rodziną wysoce zachowawczych, niekodujących sekwencji mRNA regulujących syntezę ferytyny i receptora transferyny. Występują one w regionie 5'UTR wszystkich ferytynowych mRNA i w regionie 3'UTR wszystkich mRNA dla receptorów transferyny. Dwa białka znane jako IRP1 i IRP2 (ang. *Iron Regulatory Proteins*) wiążą się z dużym powinowactwem do IRE. Gdy eukariotyczna komórka jest pozbawiona żelaza, białko IRP1 niezasocjowane z jonami Fe^{2+} wiąże się z dużym powinowactwem do IRE i całkowicie inhibuje syntezę ferytyny. Kiedy w cytoplazmie wzrasta stężenie Fe^{2+} , białko IRP1 wiążąc jony żelaza oddysocjuje od IRE aktywując proces translacji.

Element IRE mRNA dla ferrytyny ma tendencję do tworzenia struktury drugorzędowej typu spinki do włosów z wybrzuszeniem. Wszystkie znane sekwencje IRE zawierają wysoce zachowawczą 6-nukleotydową pętlę apikalną o sekwencji 5'-CAGUGX-3', gdzie X może być zarówno A, C i U, lecz nigdy G. Jak wcześniej pokazano metodami NMR (27,28), pomiędzy pierwszą i piątą zasadą wewnątrz apikalnej pętli IRE tworzy się para C-G; tym samym rozmiar pętli ulega zmniejszeniu.

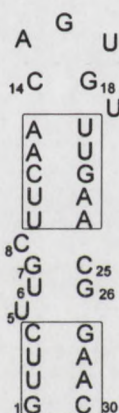
Moim zadaniem było ustalenie struktury przestrzennej badanej cząsteczki RNA, a w szczególności poznanie oddziaływań pomiędzy zasadami w regionie wybrzuszenia (wewnętrznej pętli), którego występowanie sugerowały obliczenia wykonane za pomocą programów GCG-9.0 (Wisconsin) czy MFOLD (29).

Wybrzuszenia (ang. *bulge*, *internal loops*) stanowią ważną klasę funkcjonalnych miejsc w strukturach RNA. Przykładem cząsteczek RNA, które posiadają niezbędne dla zachowania ich biologicznej funkcji wybrzuszenia są np. elementy TAR i RRE wirusa HIV (30,31), pętla E rybosomalnego 5S RNA (32) czy pętla α -sarcyny rybosomalnego 23S RNA (33).

A.)



B.)



Struktury drugorzędowe IRE mRNA dla ferrytyny:

A.) obliczona, B.) otrzymana na podstawie danych NMR

W chwili podjęcia moich badań w 1996 roku złożoność 30-meru IRE wyznaczała granice stosowalności biomolekularnego NMR na polu badań RNA. Dysponowałam zarówno preparatami w pełni znakowanymi izotopami ^{13}C i ^{15}N (ang. *uniform labeling*) i nieznakowanymi otrzymanymi metodą *in vitro* z zastosowaniem polimerazy T7 RNA. Badania rozpocząłam od zweryfikowania oddziaływań zasada-zasada w obrębie całej cząsteczki IRE. Na podstawie analizy widm NMR stwierdziłam, że poza wskazanymi na rysunku (B.) regionami helikalnych trzonów, w strukturze IRE występują jeszcze trzy dodatkowe pary zasad (III). Potwierdziłam, że jedną z nich była para C14-G18 stabilizująca strukturę pętli apikalnej. Analiza widm NOESY oraz wartości przesunięć chemicznych atomów azotu otrzymane z widm ^1H - ^{15}N HSQC wykonanych dla cząsteczki całkowicie znakowanej izotopami $^{13}\text{C}/^{15}\text{N}$ wskazywały, że następną z dodatkowych par

A.)

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Substancje organiczne...
A) obliczenia...

Wzrost...

granice...
niezależność...
Badania...
eksperyment...
rysunek...
dotyczy...
stabilność...
chemiczne...
całkowite...

zasad jest również para G-C o geometrii Watsona-Cricka. Chociaż w widmie NOESY nie zaobserwowałam sygnałów korelacyjnych odpowiedzialnych za oddziaływania warstwowe, to jedyną dodatkową parą G-C w strukturze IRE mogła być para G7-C25 w wewnętrznej pętli.

Przesunięcia chemiczne atomów ^{15}N dwóch innych sygnałów widocznych w widmie ^1H - ^{15}N HSQC (140.47 ppm dla G-N1 i 155.30 ppm dla U-N3) sugerowały, że trzecią z powstających par zasad jest para U-G. Brakowało jednak bezpośredniego dowodu z widma NOESY gdyż nie pojawiły się charakterystyczne dla par typu *wobble* U-G sygnały korelacyjne pomiędzy dwoma protonami iminowymi (po jednym z U i G). Oczekiwane sygnały korelacyjne o silnej intensywności zaobserwowałam dopiero po obniżeniu pH roztworu z 6.8 do 5.1. Miałam w tym przypadku do czynienia z interesującym przykładem wpływu warunków środowiska na oddziaływanie zasada-zasada. Obniżenie pH roztworu spowodowało poszerzenie niektórych sygnałów rezonansowych oraz dodatkowo pojawienie się nowych. W oparciu o wnikliwą analizę widm NMR, sygnały te przypisałam strukturze, w której zamiast pary U6-G26 zawiązuje się para U5-G26. Okazało się, że rozkład statystyczny konformerów RNA pozostających ze sobą w równowadze zależny jest od pH. W wyższym pH dominuje struktura z parą U6-G26. Protony iminowe biorące udział w tworzeniu pary U-G ulegają bardzo znacznej wymianie z wodą. Szybkość tej wymiany również zależna jest od pH.

W tym miejscu chciałabym dodać, że otrzymana przeze mnie struktura drugorzędowa cząsteczki IRE różni się od tej sugerowanej przez obliczenia. Dopiero najnowsza wersja programu MFOLD 3.1 (34,35) wykazuje tworzenie się pary G7-C25. W obliczonej strukturze proponowane jest jednak zawiązywanie się pary U5-G26, zamiast U6-G26. Pomimo udoskonalenia parametrów termodynamicznych tego programu, nadal nie jest przewidywane tworzenie wewnątrz apikalnej pętli pary C14-G18, której istnienie zostało potwierdzone metodami NMR (27,28,III).

Informacje otrzymane drogą analizy widm wykonanych zarówno dla cząsteczki całkowicie znakowanej izotopami $^{13}\text{C}/^{15}\text{N}$ jak i dla cząsteczki o naturalnej zawartości izotopów posłużyły mi do zbudowania modelu elementu IRE za pomocą programu MCSYM (*Macromolecular Conformation by SYMBOLIC computation*) (36,37). MC-SYM jest algorytmem obliczeniowym dedykowanym dla modelowania trójwymiarowych struktur RNA [np. tRNA (38) czy aptamerów RNA (39)]. Wykorzystuje on istniejącą eksperymentalną bazę danych wszystkich konformacji reszt nukleotydowych

zamiast jest to...

zakończ...

wartości...

wzrost...

Przed...

widnia...

traci...

z widnia...

zyski...

Określenie...

obniżenie...

interakcje...

zasada...

rozważ...

widm...

się...

ze sobą...

G26, P...

wymiar...

W...

drug...

najważ...

oblicz...

U6-G26...

nie jest...

zostało...

Inform...

całkow...

izotop...

MCSYM...

jest alg...

strukt...

ekspery...

(krystalografia, NMR) w połączeniu z otrzymanymi dla badanej cząsteczki więzami z eksperymentów NMR. W obliczeniach zdecydowałam się użyć więzy strukturalne uzyskane dla pH 6.8, a więc warunków bardziej zbliżonych do naturalnych.

Trzon modelu cząsteczki IRE stanowią dwie helisy typu A-RNA rozdzielone dynamicznym regionem wewnętrznego wybrzuszenia. Zawijające się w wewnętrznej pętli dwie pary zasad U6-G26 i G7-C25 przyczyniają się do częściowego ustrukturalizowania tego regionu. Podczas gdy reszta U5 w większości wygenerowanych przez mnie struktur jest schowana do wnętrza helisy, zachowawcza reszta C8 wystaje na zewnątrz i dynamika sprawia, że jej pozycja jest słabo zdefiniowana. Podobnie jak w innych pracach (27,28,40), najmniej informacji uzyskałam odnośnie konformacji apikalnej pętli spinki IRE. Poza parą zasad C14-G18 pozostałe reszty cukrowe w pętli miały znaczny procent konformacji C-2'-endo.

W tym miejscu chciałabym zaznaczyć, że w czasie gdy publikacja o strukturze badanego przeze mnie fragmentu IRE (III) była zaakceptowana do druku, ukazało się doniesienie z laboratorium Pardi'ego (40) dotyczące badań NMR struktury cząsteczki IRE o innej sekwencji. Badany przez tę grupę 29-nukleotydowy fragment RNA posiadał zmodyfikowaną, dla poprawienia wydajności transkrypcji, sekwencję elementu IRE znajdującego się w mRNA łańcucha H ludzkiej ferytyny. Struktura drugorzędowa tego fragmentu to trzon o długości 11 par zasad z mononukleotydowym wybrzuszeniem (zachowawcza reszta cytozyny) występującym w odległości pięciu par zasad od terminalnej pętli sześci nukleotydowej.

Chciałabym podkreślić, że otrzymany przez mnie model cząsteczki IRE (III) jest zbliżony do struktury NMR opublikowanej przez Pardi'ego (40). W obu przypadkach mamy do czynienia z dobrze zdefiniowanymi krótkimi fragmentami helikalnymi, jednakże wzajemna orientacja osi głównych tych helis nie jest ściśle ustalona. W strukturze proponowanej przez Pardi'ego pary zasad w górnym trzonie są słabiej zdefiniowane, szczególnie w pobliżu wybrzuszenia i w sąsiedztwie pętli apikalnej.

Pary G-U są najczęściej spotykanym typem niekanonicznych oddziaływań zasada-zasada w RNA. Miejsca występowania par G-U nie są przypadkowo rozmieszczone w sekwencji RNA, lecz występują na ogół w funkcjonalnie ważnych regionach, m.in. wybrzuszeniach (41). Rozpoczynając moje badania nie podejrzewałam, że zależne od pH formowanie drugorzędowej struktury IRE, a szczególnie tworzenie dynamicznej pary G-U

(fizjologia, 1978)
obserwacje
wynikami do
Tymczasowo
dynamizm
pół-
struktur
przez
zwym
innych
pół-
proces
W
badania
dotyczy
o-
zmodyfik
miejsc
fragment
(zachowa
tymczas
Charakter
zdoby
mamy
wzajem
propoz
szereg
Tym
zasada
skwar
wymiar
formo

w regionie wybrzuszenia, jest związane także z procesem specyficznego kompleksowania jonów magnezu w tym regionie.

Aby cząsteczka RNA, która jest polianionem, mogła utworzyć funkcjonalną biologicznie strukturę trzeciorzędową niezbędna jest obecność kationów. Wszystkie znane, silnie pofałdowane, cząsteczki RNA zawiązują struktury drugorzędowe przy braku jonów metalu, natomiast nie wiemy czy w sposób optymalny; niewiele też tworzy się elementów ich struktury trzeciorzędowej. Podczas gdy większość kationów jest zdelokalizowana lub kompensuje reszty fosforanowe, niektóre z nich zajmują specyficzne miejsca w strukturze RNA (ang. *magnesium pockets*) kształtowane przez odpowiednie zwinięcie się łańcucha oligorybonukleotydowego. Jak wykazują badania krystalograficzne (42-44), specyficzne wiązanie hydratowanych jonów metalu, głównie magnezu, do RNA może następować albo bezpośrednio (ang. *inner sphere complexes*) albo poprzez ich sferę hydratacyjną (ang. *outer sphere complexes*). Oddziaływania jonów magnezu, w które zaangażowane są zasady, następują na ogół poprzez sferę hydratacyjną jonów Mg^{2+} . W oddziaływaniach tego typu najczęściej biorą udział atomy azotu N7 puryn oraz atomy tlenu grup karbonylowych (45,46).

Niestety, metody NMR stosowane do określania struktur RNA w roztworze nie dostarczają dokładnych informacji o miejscu wiązania jonów Mg^{2+} . Zmiany obserwowane w widmie NMR cząsteczek RNA pod wpływem Mg^{2+} trudno rozdzielić pomiędzy te spowodowane bezpośrednim oddziaływaniem z jonami od tych, które są wynikiem zmian w strukturze przestrzennej badanej cząsteczki. Ponadto obecność jonów metalu może być czynnikiem przyczyniającym się do degradacji badanego obiektu. Z tych względów bardzo rzadko przeprowadza się eksperymenty NMR miareczkowania magnezem, gdyż potrzeba dla nich miligramowych ilości RNA.

Z uwagi na porównywalną wielkość promieni jonowych oraz geometryczne podobieństwo pomiędzy hydratowanymi jonami magnezu, $Mg(H_2O)_6^{2+}$, a jonom heksaaminakobaltu(III), $Co(NH_3)_6^{3+}$, zaproponowano jego zastosowanie jako sondy dla zlokalizowania miejsc specyficznego wiązania jonów magnezu do cząsteczek RNA (47,48). Użycie do badań w roztworze jonu $Co(NH_3)_6^{3+}$ zamiast $Mg(H_2O)_6^{2+}$ ma tę przewagę, że możliwe jest obserwowanie międzycząsteczkowego efektu NOE pomiędzy protonami pochodzącymi z RNA a protonami jonu kompleksowego, a tym samym precyzyjne umiejscowienie tego jonu w strukturze kwasu nukleinowego.

w regionie wybranych...
jony magnezu...

Abi...
biologicznych...

silnie...
metale...

ich...
kompleksy...

RNA (arg...
oligomerizacji...

wystąpi...
bezpłodnie...

owr...
zadby...

tego...
kubkowy...

W...
dotyczy...

w widmie...
spowodowane...

w strukturze...
czynniki...

zadko...
dla nich...

X...
podobnie...

bezsamochodny...
zlokalizowana...

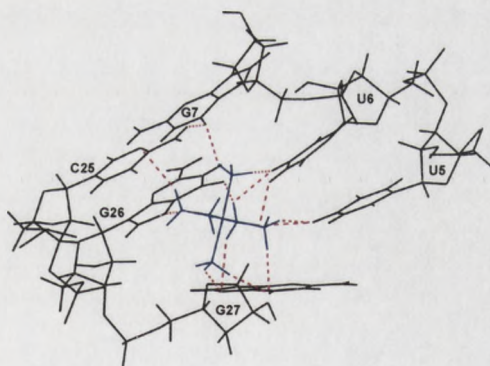
(47A2)...
przeważ...

protonni...
pewny...

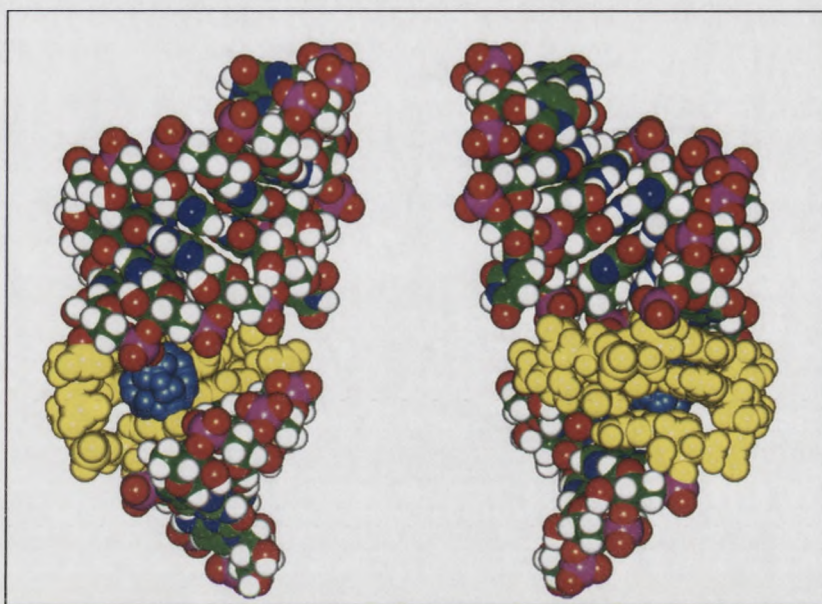
Metodę tę zastosowałam dla elementu IRE (III). Zaobserwowałam wzrost temperatury topnienia T_m 30-meru IRE RNA w obecności $\text{Co}(\text{NH}_3)_6^{3+}$ o 4°C w przypadku stechiometrii 1:1. Gdy stechiometria RNA : jon metalu zwiększyła się do 1:2 i 1:3, wzrosła również stabilność RNA odpowiednio o 8°C i 10°C . Również zmiany zachodzące w widmie ^1H NMR pod wpływem dodawanego $\text{Co}(\text{NH}_3)_6^{3+}$ wskazywały, że różne regiony cząsteczki IRE odmiennie reagują na obecność tego czynnika. Dodatek $\text{Co}(\text{NH}_3)_6^{3+}$ spowodował największe zmiany przesunięć chemicznych protonów znajdujących się w obrębie wybrzuszenia.

W celu sprawdzenia czy zaproponowany przeze mnie model cząsteczki zawiera kieszeń wiązania jonu metalu i koresponduje z rezultatami miareczkowania $\text{Co}(\text{NH}_3)_6^{3+}$, przeprowadziłam symulację komputerową opartą o protokół DOCKING (analiza potencjału elektrostatycznego i oddziaływań Van der Waalsa) zmierzającą do znalezienia potencjalnych miejsc wiązania $\text{Co}(\text{NH}_3)_6^{3+}$ w wygenerowanym przez mnie modelu cząsteczki (III). We wszystkich otrzymanych z symulacji kompleksach IRE/ $\text{Co}(\text{NH}_3)_6^{3+}$, jon metalu wiąże się w dużej bruzdzie w regionie wewnętrznej pętli (patrz rysunki), gdzie obecność elektrojemnych atomów zasad przyczynia się do specyficznego wiązania jonu metalu, a pole elektrostatyczne regionu helikalnego zaburzone jest przez obecność niekanonicznych oddziaływań.

W generowanych strukturach kompleksu RNA : jon metalu, jony kompleksowe $\text{Co}(\text{NH}_3)_6^{3+}$ zaangażowane są w sieć wiązań wodorowych. Struktury kompleksów grupują się w dwóch klasach. W pierwszej wiązania wodorowe występują pomiędzy $\text{Co}(\text{NH}_3)_6^{3+}$ a resztami G27, G26, U5, U6 i C25; w drugiej (patrz rysunek) dodatkowo występuje wiązanie wodorowe do G7.

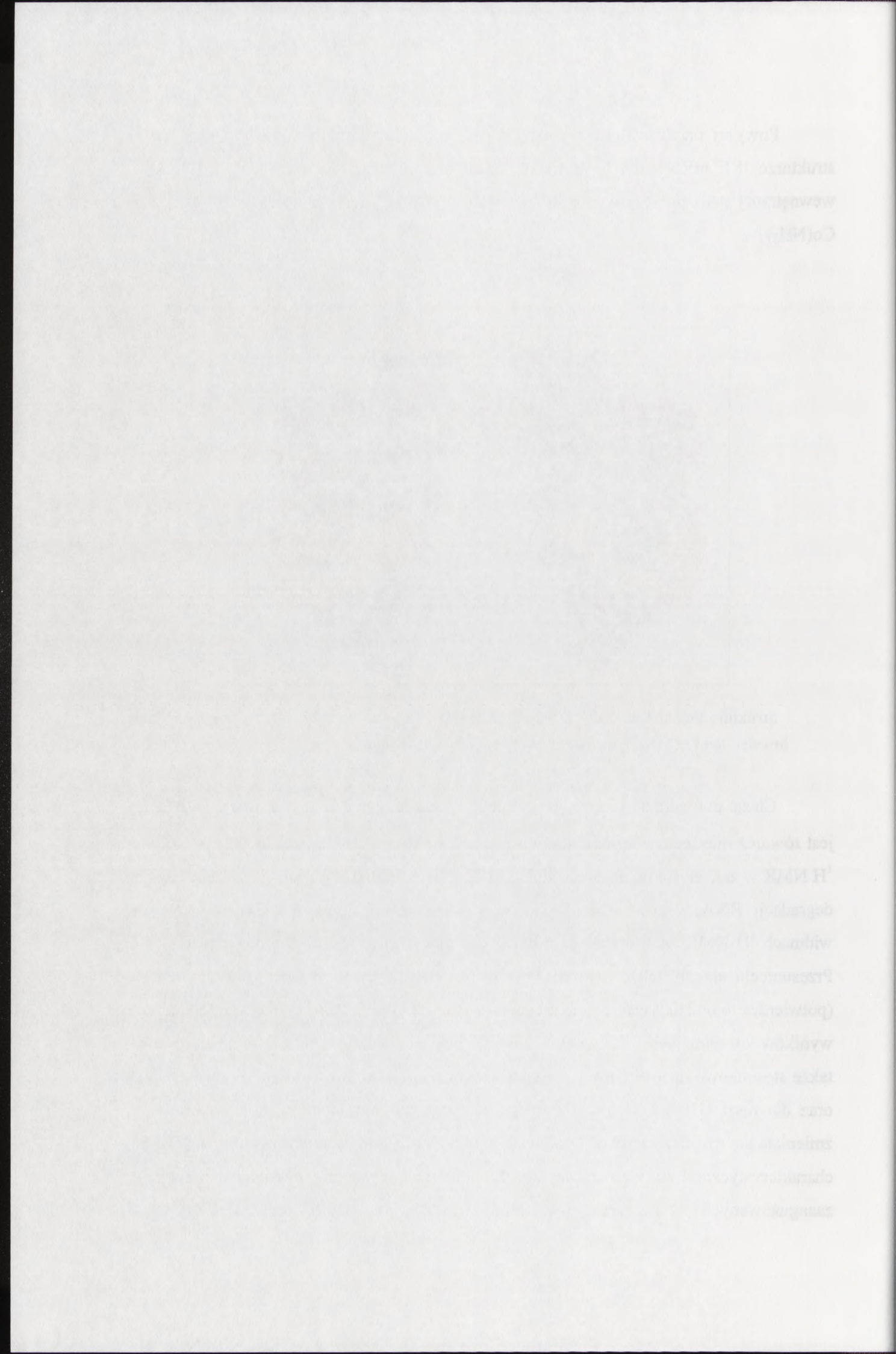


Powyżej przedstawione wyniki pozwoliły mi na sformułowanie hipotezy, że w strukturze IRE mRNA dla ferrytyny obecność nieklasycznej pary U6-G26 (pH 6.8) w wewnętrznej pętli przyczynia się do powstania polarnej kieszeni specyficznie wiążącej $\text{Co}(\text{NH}_3)_6^{3+}$.



Struktura kompleksu IRE/ $\text{Co}(\text{NH}_3)_6^{3+}$. Z lewej – widok na dużą bruzdę, z prawej - na małą bruzdę. Jon $\text{Co}(\text{NH}_3)_6^{3+}$ zaznaczony kolorem jasnoniebieskim.

Chcąc potwierdzić eksperymentalnie, że kieszeń specyficznie wiążąca $\text{Co}(\text{NH}_3)_6^{3+}$ jest również miejscem wiązania jonów magnezu śledziłam zmiany zachodzące w widmie ^1H NMR w zależności od stężenia MgCl_2 (IV). Miałam dużo szczęścia – nie zauważyłam degradacji RNA w tych warunkach. Największe zmiany przesunięć chemicznych w widmach ^1H NMR zaobserwowałam dla reszty guanozyny G26 w obrębie pary U6-G26. Przesunięciu ulegały także położenia sygnałów pozostałych reszt pętli wewnętrznej: U6 (potwierdza to oddziaływanie jonu magnezu z parą U6-G26), G27 i G7. W odróżnieniu od wyników otrzymanych z eksperymentów z heksaaminakobaltem(III), zarejestrowałam także stosunkowo duże zmiany wywołane obecnością jonów Mg^{2+} dla protonu reszty G18 oraz dla reszt U10 i U21 z górnego trzonu. Wraz ze wzrostem stężenia jonów Mg^{2+} zmieniała się również szerokość połówkowa niektórych linii rezonansowych. Szczególnie charakterystyczne było wyostrenie się sygnałów pochodzących od protonów iminowych zaangażowanych w tworzenie par zasad wewnątrz obu pętli, terminalnej (G18) i



wewnętrznej (G7). Obserwowane zmiany szerokości linii sugerują, że obecność jonów Mg^{2+} przyczynia się do wzrostu stabilności regionów pętlowych w strukturze IRE. Ilościowe zmiany obserwowane w widmie NMR w obecności jonów Mg^{2+} potwierdzają, że region wewnętrznej pętli jest również miejscem wiązania Mg^{2+} .

Uzyskane wyniki strukturalne (III,IV) stały się bodźcem i dały podstawy do szerszych badań NMR dotyczących oddziaływania RNA z jonami Mg^{2+} . Korzystając z opracowanej ostatnio w naszej Pracowni chemicznej metody znakowania izotopami ^{13}C wybranych reszt nukleotydowych drogą totalnej syntezy chemicznej RNA (8) rozwijamy badania dotyczące specyficznych oddziaływań Mg^{2+} z 29-merem RNA o sekwencji elementu TAR HIV-1.

Województwo Lubelskie

Urząd Marszałkowski Województwa Lubelskiego

Biuro Regionalne w Lublinie

ul. Świdnicka 10, 20-030 Lublin

tel. 81 425 21 00, 81 425 21 01

fax 81 425 21 02

www.lublin.gov.pl

Biuro Regionalne w Lublinie

ul. Świdnicka 10, 20-030 Lublin

tel. 81 425 21 00, 81 425 21 01

fax 81 425 21 02

www.lublin.gov.pl

4. WPLYW N⁶-(N-TREONYLOKARBONYLO)ADENOZYNY (t⁶A) NA ODDZIAŁYWANIA ZASADA-ZASADA W STRUKTURZE 17-MERU RNA, ANALOGU RAMIENIA ANTYKODONOWEGO tRNA^{Lys3}.

Transferowe kwasy rybonukleinowe należą do najbardziej złożonych i egzotycznych struktur RNA. Głównym wyróżnikiem pomiędzy tRNA a innymi klasami cząsteczek RNA jest duża liczba i różnorodność modyfikowanych nukleozydów obecnych w tRNA. Z poznanych do tej chwili około 100 różnych modyfikacji, wszystkie są wprowadzane posttranskrypcyjnie. Dotychczasowa wiedza dotycząca budowy tRNA opiera się o stosunkowo wcześnie poznane struktury kryształów m.in. tRNA^{Phe} z drożdży (1974) (49,50). Złożoność i wielkość cząsteczek tRNA (około 80 reszt nukleotydowych) uniemożliwia analizę ich struktury metodami NMR. Z tych względów badaniami NMR objęto określone domeny architektury tRNA, najczęściej ramiona tej cząsteczki, które klasyfikowane są jako spinki RNA.

Jedyną drogą dostępu do modyfikowanych i hipermodyfikowanych fragmentów tRNA jest ich totalna synteza chemiczna. W 1978 roku opracowano w naszym laboratorium po raz pierwszy syntezę pętli antykodonowej inicjatorowego tRNA o sekwencji CCCAU(t⁶A)A, zawierającej hipermodyfikowany nukleozyd N⁶-(N-treonylokarbonylo)adenozynę (t⁶A) (51). W owym czasie synteza 17-meru o sekwencji pełnego ramienia antykodonowego okazała się niemożliwa. Syntezę 17-merów zawierających tę modyfikację o sekwencjach tRNA^{Met} CAGGGCUCAU(t⁶A)ACCCUG i tRNA^{Lys3} UCAGACU(mnm5s2U)UU(t⁶A)A(Ψ)CUGA opracowano dopiero w ubiegłym roku w dwóch laboratoriach (52,53).

Ciągle rosnące zagrożenie epidemią AIDS spotęgowało zainteresowanie strukturą tRNA^{Lys3}. Oprócz typowej dla wszystkich cząsteczek tRNA funkcji w procesie syntezy białka, ludzki tRNA^{Lys3} jest również starterem (ang. *primer*) dla odwrotnej transkryptazy HIV i odgrywa kluczową rolę w cyklu życiowym retrowirusa HIV-1. Dwie posttranskrypcyjne modyfikacje wyróżniają tRNA^{Lys3} ludzki i z *Escherichia coli*. W ludzkim tRNA^{Lys3} w pozycji 34 znajduje się 5-acetylo-2-tiourydyna (*mcm*⁵*s*²*U*) a w pozycji 37, przyległej do antykodonu od końca 3' jest zawsze obecna 2-metylotio-N⁶-(N-treonylokarbonylo)adenozyna (*ms*²*t*⁶*A*). W tRNA^{Lys3} z *Escherichia coli* występują kolejno w tym miejscu 5-metyloaminometylo-2-tiourydyna (*mnm*⁵*s*²*U*) oraz t⁶A. Zarówno oddziaływania pomiędzy pętlą antykodonu cząsteczki tRNA^{Lys3} i bogatą w reszty adeniny

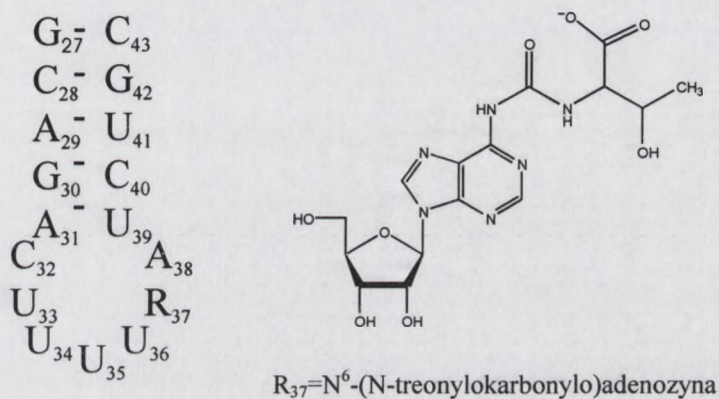
Transkrypcja jest procesem, w którym informacja genetyczna zapisana w DNA jest kopiowana do RNA. Głównym produktem transkrypcji jest mRNA, które następnie jest przetworzone i przetransportowane do cytoplazmy, gdzie służy jako szablon do syntezy białek. Istotnym elementem transkrypcji jest cząsteczka RNA polimerazy, która katalizuje łączenie nukleotydów. W RNA 5' pozostaje do tej pory wolny koniec, który umożliwia przyłączenie cząsteczki kapek 5' mCPPP, co zapobiega degradacji RNA. W RNA 3' znajduje się koniec poli(A), który jest dodany przez enzymy w procesie dojrzewania mRNA. Istotnym elementem transkrypcji jest również cząsteczka RNA polimerazy, która katalizuje łączenie nukleotydów. W RNA 5' pozostaje do tej pory wolny koniec, który umożliwia przyłączenie cząsteczki kapek 5' mCPPP, co zapobiega degradacji RNA. W RNA 3' znajduje się koniec poli(A), który jest dodany przez enzymy w procesie dojrzewania mRNA.

Jedną z funkcji RNA jest transport aminokwasów do rybosomu, gdzie następuje synteza białka. Istotnym elementem transkrypcji jest cząsteczka RNA polimerazy, która katalizuje łączenie nukleotydów. W RNA 5' pozostaje do tej pory wolny koniec, który umożliwia przyłączenie cząsteczki kapek 5' mCPPP, co zapobiega degradacji RNA. W RNA 3' znajduje się koniec poli(A), który jest dodany przez enzymy w procesie dojrzewania mRNA. Istotnym elementem transkrypcji jest również cząsteczka RNA polimerazy, która katalizuje łączenie nukleotydów. W RNA 5' pozostaje do tej pory wolny koniec, który umożliwia przyłączenie cząsteczki kapek 5' mCPPP, co zapobiega degradacji RNA. W RNA 3' znajduje się koniec poli(A), który jest dodany przez enzymy w procesie dojrzewania mRNA.

Ciągle rozwija się wiedza na temat roli RNA w procesach biologicznych. Istotnym elementem transkrypcji jest cząsteczka RNA polimerazy, która katalizuje łączenie nukleotydów. W RNA 5' pozostaje do tej pory wolny koniec, który umożliwia przyłączenie cząsteczki kapek 5' mCPPP, co zapobiega degradacji RNA. W RNA 3' znajduje się koniec poli(A), który jest dodany przez enzymy w procesie dojrzewania mRNA. Istotnym elementem transkrypcji jest również cząsteczka RNA polimerazy, która katalizuje łączenie nukleotydów. W RNA 5' pozostaje do tej pory wolny koniec, który umożliwia przyłączenie cząsteczki kapek 5' mCPPP, co zapobiega degradacji RNA. W RNA 3' znajduje się koniec poli(A), który jest dodany przez enzymy w procesie dojrzewania mRNA.

pętlą HIV-1 RNA, jak i oddziaływania kompleksu tRNA^{Lys3}/HIV-1 z odwrotną transkryptazą zależą silnie od modyfikacji (54).

Znając nasze zainteresowanie rolą hipermodyfikacji t⁶A w tRNA, prof. P.F.Agris (NCSU, Raleigh, USA) zwrócił się do mnie z prośbą o pilną pomoc w rozwiązaniu struktury w roztworze 17-meru RNA, analogu ramienia antykodonowego tRNA^{Lys3} z *E.coli*. Przedmiotem moich badań miały być dwie cząsteczki: zawierająca hipermodyfikowany nukleozyd t⁶A w pozycji 37 (ASL^{Lys3}-t⁶A) oraz niemodyfikowana cząsteczka ASL^{Lys3}-A.



Sekwencja ramienia antykodonowego ASL^{Lys3}-t⁶A

17-nukleotydomy analog ramienia antykodonowego cząsteczki tRNA^{Lys3} odpowiadający sekwencji ASL (ang. - *Anticodon Stem Loop*) posłużył mi ponadto jako model do zbadania wpływu modyfikacji zasad, w tym przypadku adeniny, na strukturę i oddziaływania zasada-zasada w regionie pętlowym spinek RNA (V). Z badań biochemicznych wiadomo, że zarówno cząsteczka tRNA^{Lys3} jak i 17-nukleotydomy fragment ASL^{Lys3}-A pozbawione modyfikacji nie wiążą się do odpowiednio zaprogramowanego rybosomu (55). Dla przywrócenia zdolności wiązania wystarczająca jest już obecność jednej z modyfikacji, *mnm*⁵*s*²*U* lub t⁶A, jednak ich powinowactwo jest mniejsze niż w przypadku obu modyfikacji lub natywnego tRNA^{Lys3} (56). Powyższe obserwacje wskazywały, że obecność lub brak modyfikowanych nukleotydów w tej cząsteczce może prowadzić do dużych zmian konformacyjnych w obrębie pętli.

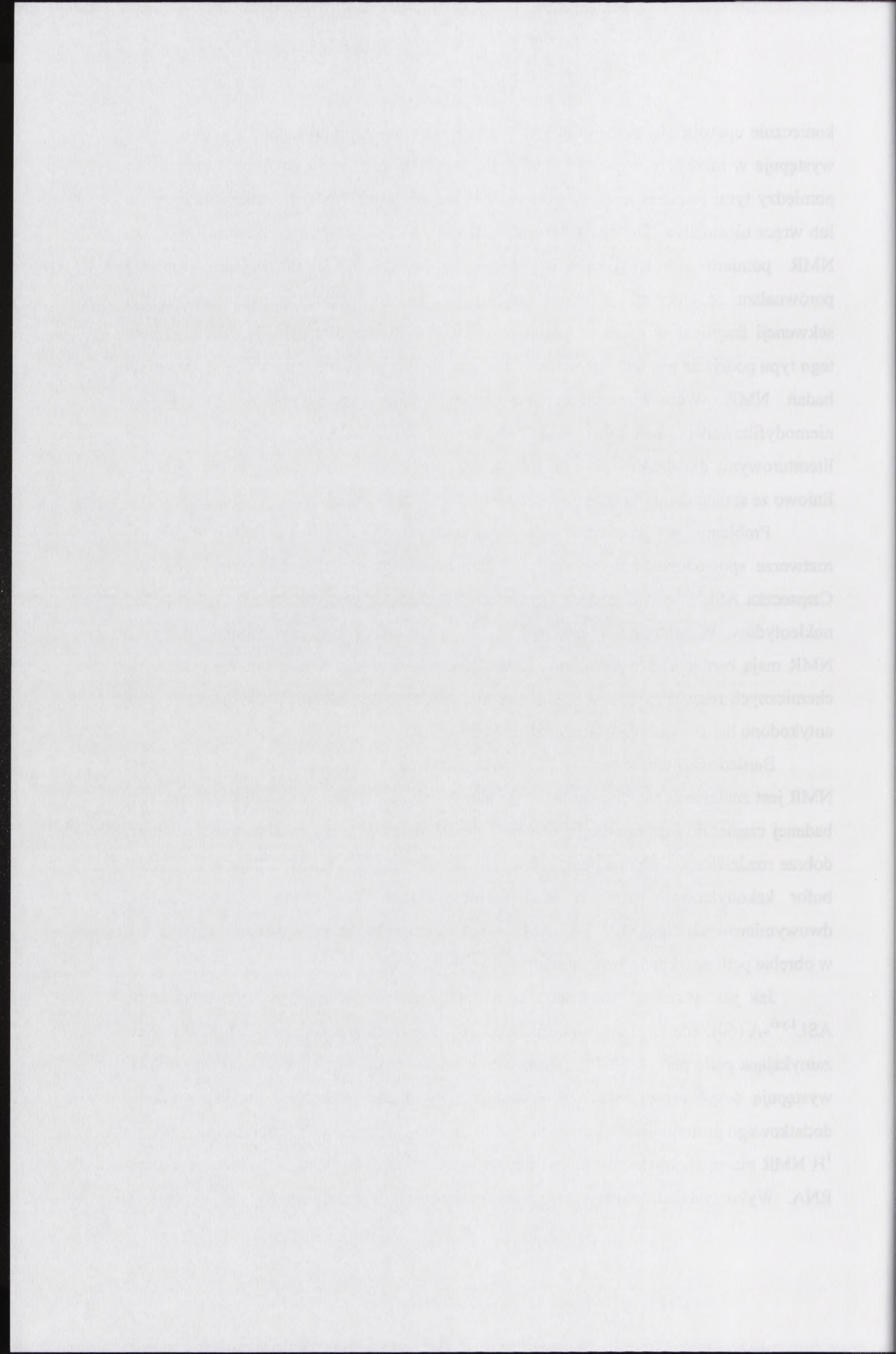
Analiza konformacyjna pętli w strukturach spinkowych, ASL^{Lys3}-A i ASL^{Lys3}-t⁶A₃₇, zawierających dużą liczbę niesparowanych nukleotydów okazała się niezwykle złożona. Przed przystąpieniem do badań struktur spinkowych RNA metodami NMR należy

koniecznie upewnić się, że dany oligomer rzeczywiście przyjmuje żadaną strukturę, a nie występuje w roztworze w postaci wzajemnie komplementarnego dupleksu. Rozróżnienie pomiędzy tymi formami na podstawie standardowych widm NMR jest niezwykle trudne lub wręcz niemożliwe. Dlatego w swoich badaniach wykorzystałam gradientową technikę NMR pomiaru translacyjnego współczynnika dyfuzji (57). Otrzymane wartości porównałam ze znanymi (i pomierzonymi) wartościami współczynników dla innych sekwencji fragmentów kwasów nukleinowych o różnej długości (58,59). Wielką zaletą tego typu podejścia jest fakt, że pomiaru dokonuje się bezpośrednio na próbce używanej do badań NMR. Wartości pomierzonych przez mnie współczynników dyfuzji dla niemodyfikowanej cząsteczki $ASL^{Lys3}\text{-A}$ i dla $ASL^{Lys3}\text{-t}^6\text{A}$ są zgodne z danymi literaturowymi dla cząsteczek o takim ciężarze cząsteczkowym (58) i są skorelowane liniowo ze standardami dla heksameru, dodekameru i 28-meru RNA.

Problemy w określeniu cech strukturalnych $ASL^{Lys3}\text{-A}$ i $ASL^{Lys3}\text{-t}^6\text{A37}$ w roztworze spowodowane są obecnością licznych konformacji lub znaczącej dynamiki. Cząsteczka $ASL^{Lys3}\text{-t}^6\text{A37}$ posiada typową dla antykodonu pętlę składającą się z siedmiu nukleotydów. W centralnej części pętli znajdują się cztery urydyny, które w widmach NMR mają bardzo zbliżone wartości przesunięć chemicznych. Podobieństwo przesunięć chemicznych reszt urydynowych, które są niezwykle istotne dla struktury badanych pętli antykodonu bardzo komplikowało analizę widm NMR.

Bardzo ważnym etapem w badaniach polianionowych cząsteczek RNA metodami NMR jest znalezienie takich warunków pomiarowych, w których labilność konformacyjna badanej cząsteczki jest silnie ograniczona co sprawia, że sygnały rezonansowe są ostre i dobrze rozdzielone. Dobrane przez mnie warunki dla $ASL^{Lys3}\text{-A}$ i $ASL^{Lys3}\text{-t}^6\text{A}$ to: 10mM bufor kakodylanowy, pH 5.6 oraz 0.1mM EDTA. Porównanie widm jedno- i dwuwymiarowych dla $ASL^{Lys3}\text{-A}$ i $ASL^{Lys3}\text{-t}^6\text{A}$ wskazywało na duże zamiany strukturalne w obrębie pętli antykodonowej indukowane modyfikacją t^6A .

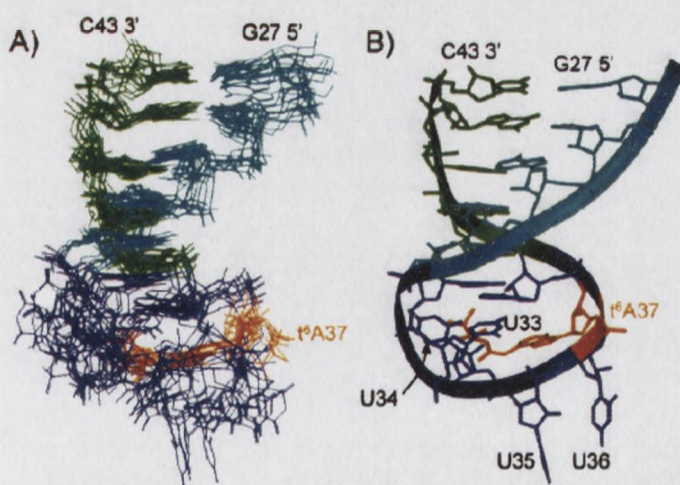
Jak już wcześniej pokazano dla niemodyfikowanego analogu pętli antykodonu $ASL^{Lys3}\text{-A}$ (60), adenozyne w pozycji 38 (A38) ulega protonacji przy pH 5 i tworzy się zamykająca pętlę para $A^+38\text{-C32}$. Zarówno w spinkach DNA jak i RNA pary typu $A^+\text{-C}$ występują dosyć często i na ogół sąsiadują z regionem helikalnym. Szybka wymiana dodatkowego protonu nieklasycznej pary $A^+\text{-C}$ z wodą jest typowa i z tego powodu widma ^1H NMR nie są zazwyczaj użyteczne dla stwierdzenia ich obecności w badanej strukturze RNA. Wykorzystałam dlatego zależność przesunięcia chemicznego atomu węgla C2



adenozyny od pH jako metodę detekcji jej formy protonowanej (61,62.). Wykonane przeze mnie dla trzech różnych wartościach pH (5.2, 6.0 i 7.2) widma ^1H - ^{13}C HSQC pokazały, że wprowadzenie modyfikacji $t^6\text{A}37$ nie wpływa na zmianę protonacji adenozyiny A38 w stosunku do analogicznej reszty adeniny w cząsteczce niemodyfikowanej $\text{ASL}^{\text{Lys}3}$. W pH 5.6 adenozyina A38 jest protonowana w pozycji N1 dla obu cząsteczek, $\text{ASL}^{\text{Lys}3}\text{-A}$ i $\text{ASL}^{\text{Lys}3}\text{-}t^6\text{A}$.

We wcześniejszych badaniach cząsteczki $\text{ASL}^{\text{Lys}3}\text{-A}$ pozbawionej modyfikacji (60), oprócz zawiązywania się pary $\text{A}^+\text{-C}$, postulowano także obecność dodatkowej, słabej pary U33-A37 wewnątrz pętli. W dobranych przeze mnie warunkach roztworu, porównując widma ^1H NMR dla $\text{ASL}^{\text{Lys}3}\text{-A}$ i $\text{ASL}^{\text{Lys}3}\text{-}t^6\text{A}$, łatwo było zauważyć, że sygnał pochodzący od protonu iminowego reszty U33 w widmie cząsteczki niemodyfikowanej znika w obecności $t^6\text{A}37$. Tym samym obecność modyfikacji w pozycji 37 zapobiega zawiązywaniu pary U33- $t^6\text{A}37$.

Zastosowanie różnych homo- i heterojądrowych technik NMR szczegółowo opisanych w pracy (V) (NOESY, TOCSY, DQF-COSY, HSQC, ^1H - ^{31}P -HETCOR, hetero-TOCSY-NOESY) pozwoliło mi na przypisanie większości sygnałów ^1H i ^{31}P . Analiza ilościowa widm NOESY wykonanych dla $\text{ASL}^{\text{Lys}3}\text{-}t^6\text{A}37$ umożliwiła zgromadzenie 220 eksperymentalnych więzów odległościowych, a 46 więzów kątowych zostało otrzymanych z widm DQF-COSY. Więzy te zostały następnie wykorzystane do obliczenia struktury cząsteczki metodami dynamiki molekularnej (protokół DG – *Distance Geometry*, SA-*Simulated Annealing*, pole siłowe AMBER – pakiet oprogramowania MSI).

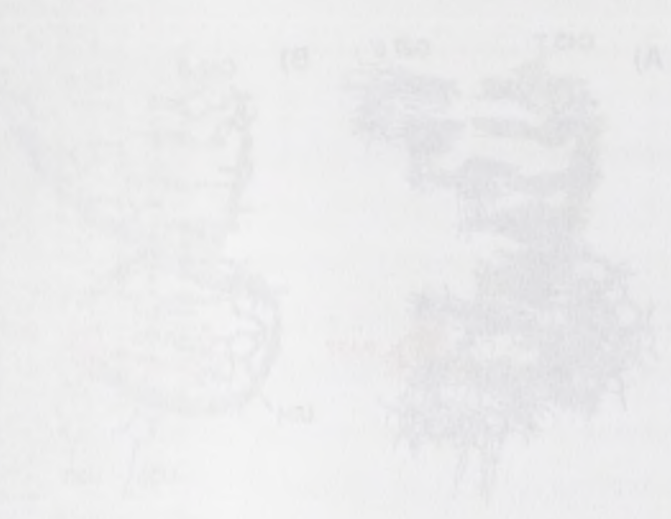


Struktura $\text{ASL}^{\text{Lys}3}\text{-}t^6\text{A}37$ otrzymana na podstawie badań NMR. Reszta $t^6\text{A}37$ zaznaczona na pomarańczowo. A) - nałożenie 10 struktur o najmniejszej energii, B) – struktura uśredniona.

atmosfery od pH jako metody detekcji jej obecności. W tym celu w celu wyznaczenia punktu izoelektrycznego (pI) białka, w którym nie występuje ładunek netto, wykonano pomiary w zakresie pH 2-12. Wyniki pomiarów przedstawiono na rysunku 1. Wskazuje on, że pI białka wynosi 4,5. W tym celu wykonano również pomiary w zakresie pH 2-12. Wyniki pomiarów przedstawiono na rysunku 1. Wskazuje on, że pI białka wynosi 4,5.

W celu wyznaczenia punktu izoelektrycznego (pI) białka, w którym nie występuje ładunek netto, wykonano pomiary w zakresie pH 2-12. Wyniki pomiarów przedstawiono na rysunku 1. Wskazuje on, że pI białka wynosi 4,5. W tym celu wykonano również pomiary w zakresie pH 2-12. Wyniki pomiarów przedstawiono na rysunku 1. Wskazuje on, że pI białka wynosi 4,5.

Zastosowanie różnych metod (pH, pI, pKa) pozwoliło na wyznaczenie punktu izoelektrycznego białka. Wyniki pomiarów przedstawiono na rysunku 1. Wskazuje on, że pI białka wynosi 4,5. W tym celu wykonano również pomiary w zakresie pH 2-12. Wyniki pomiarów przedstawiono na rysunku 1. Wskazuje on, że pI białka wynosi 4,5.



Rysunek 1. Struktura białka ASL¹⁴⁷ (A) - widok z przodu, (B) - widok z boku. Wskazano na rysunku punkty izoelektryczny (pI) i punkty pKa.

Konformacja pętli ASL^{Lys3} - t^6A37 różni się znacznie od opublikowanej wcześniej w laboratorium Davis'a (60) niemodyfikowanej pętli cząsteczki ASL^{Lys3} -A. Hipermodyfikowana zasada t^6A37 przyjmuje formę płaską w regionie N',N''-dwupodstawionego wiązania ureidowego poprzez utworzenie wewnątrzcząsteczkowego wiązania wodorowego pomiędzy $N\alpha$ -H reszty treoninowej z atomem azotu N1 pierścienia purynowego. Przyczynia się to do dodatkowego wzmocnienia oddziaływań warstwowych rozciągających się w kierunku 3' cząsteczki. We wszystkich strukturach reszty t^6A37 i $A38^+$ oddziałują warstwowo, a reszta adenozyiny $A38^+$ znajduje się pod resztą urydyny U39 z trzonu.

W pozycji 34 (pozycja *wobble*) naturalnego $tRNA^{Lys3}$ z *E. coli* znajduje się modyfikowana reszta nukleotydomowa mnm^5s^2U . W analizowanej przez mnie cząsteczce ASL^{Lys3} - t^6A37 reszta niemodyfikowanej urydyny U34 w pozycji *wobble* antykodonu jest znacznie przemieszczona w strukturze pętli w stosunku do analogicznej pozycji w cząsteczce ASL^{Lys3} -A. Jednakże przestrzenna bliskość reszty U34 w stosunku do reszt $A38^+$ i t^6A37 w ASL^{Lys3} - t^6A37 oznacza, że w tego typu cząsteczce tylko dwie reszty urydyny U35 i U36, a więc pozycja druga i trzecia antykodonu, byłyby dostępne dla tworzenia specyficznych oddziaływań z kodonem.

Porównanie struktury cząsteczek modyfikowanej ASL^{Lys3} - t^6A37 (V) i niemodyfikowanej ASL^{Lys3} -A (V,60) pokazuje, że wprowadzenie reszty t^6A zmienia geometrię pary $C32$ - $A38^+$. Aby pomieścić resztę treonylową, urydyna U34 jest odsunięta od t^6A37 , z kolei reszta C32 jest przemieszczona aby pomieścić U34. W efekcie tych zmian zasady w parze $C32$ - $A38^+$ nie leżą w płaszczyźnie. Otrzymane dla uśrednionej struktury kąty wiązań wodorowych $A38-N1-H1\cdots C32-O2$ (143°) i $A38-N6-H6.1\cdots C32-N3$ (155°) zmieniają się w stosunku do analogicznych kątów w cząsteczce ASL^{Lys3} -A średnio o 17° . Należy dodać, że geometria pary $C32$ - $A38^+$ otrzymana została dla struktury, dla której w obliczeniach nie nałożono żadnych więzów na wiązanie wodorowe dla reszty A38.

Analiza struktur cząsteczki niemodyfikowanej i ASL^{Lys3} - t^6A37 wskazuje na to, że obecność t^6A37 zapobiega tworzeniu zwartej struktury pętli obserwowanej w przypadku cząsteczki ASL^{Lys3} -A.

Żadne z moich danych spektralnych (1H , ^{31}P NMR) dla cząsteczki ASL^{Lys3} - t^6A37 , ani końcowa struktura nie wykazały tworzenia się typowego dla natywnej struktury $tRNA$ motywu *U-turn*. *U-turn* jest odpowiedzialny za gwałtowną zmianę kierunku ($\sim 180^\circ$)

Konformacja pełni ASL¹⁰⁰-A37 może się zmieniać od opóźnionej wczesnej w
laboratorium Davara (60) nieopóźnionej pełni opóźnionej ASL¹⁰⁰-A.
Hipermetylowana zasada CAA7 wykazuje formę płaską w regionie N¹N²-
dwupołożeniowego wiązania wodowego poprzez utworzenie nerwowoenergetycznego
wiązania wodowego pomiędzy N¹-H i C=O wzdłuż zlokalizowanego wzdłuż płaszczyzny
przecznej. Przewidywane jest, że do lokalizacji wiązania odkształcającej wartości
rozciągającej się w kierunku 3' cząsteczki. W wyjątkach niektórych testy CAA7 i
A38¹⁰⁰ oddziałują wartościowo z formą struktury A38¹⁰⁰ znajdując się pod kątem 90 stopni
U39 x rzon.

W poręgi 34 (poręgi woda) zmutowanego RNA¹⁰⁰ x A zoll znajduje się
modyfikowana testu modyfikowana wzdłuż 10. W analizowanej przez całe cząsteczki
ASL¹⁰⁰-A37 testu niezmodyfikowanej wyjątki U34 w poręgi woda są podobne jak
znacznie zmniejszona w struktury pełni w stosunku do analogicznej poręgi w
cząsteczki ASL¹⁰⁰-A. Jednakże praktycznie identyczne testy U34 w stosunku do testu
A38¹⁰⁰-A37 w ASL¹⁰⁰-A37 oznacza, że w tego typu cząsteczki tylko dwa testy
wyjątki U32 i U36 z więc poręgi 34 i 36 nie są odpowiednio były dostępne dla
tworzenia specyficznych oddziaływań z kationem.

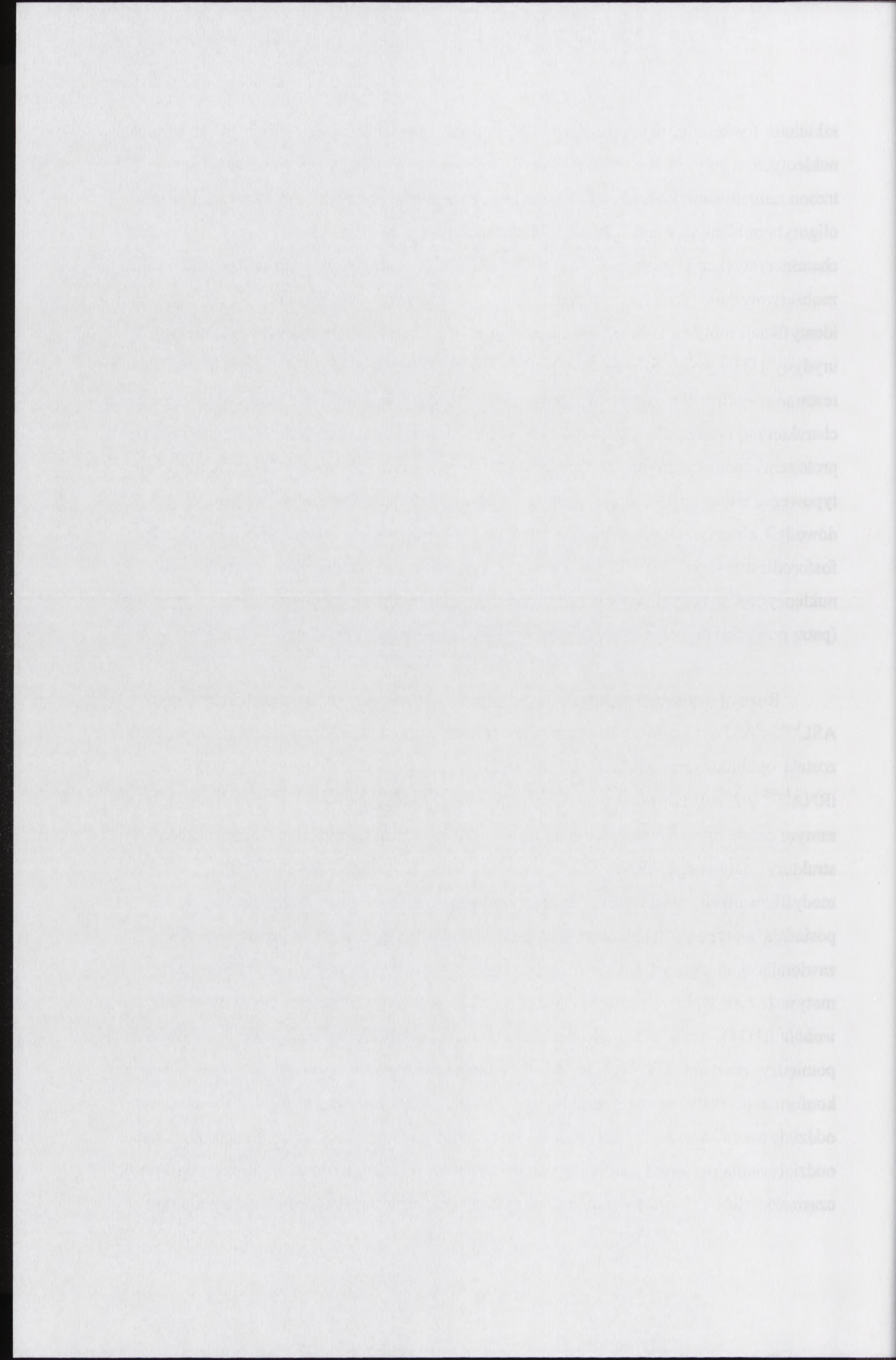
Formowanie struktury cząsteczki modyfikowanej ASL¹⁰⁰-A37 (V) i
niezmodyfikowanej ASL¹⁰⁰-A (Y,6) polega na wprowadzeniu testy A zmutacji
geometyrii par C32-A38. Aby powrócić testu trybowej wyjątki U34 jest odwrócić
od CAA7, a kolej testu C32 jest zmniejszona aby powrócić U34. W efekcie tych
zmian testy w parze C32-A38 nie jest w planyżynie. Ograniczenie dla redukcji
struktury jest wiązaniem wodowym A38-N¹H-C32-O2 (AA7) i A38-N¹H-C32-
N3 (122) zmieniają się w stosunku do analogicznych kątów w cząsteczki ASL¹⁰⁰-A
kątów o 17°. Należy dodać, że geometria par C32-A38 ograniczona została dla
struktury, dla której w obliczeniach nie należało brać pod uwagę wiązania wodowego
dla testy A38.

Analiza struktury cząsteczki niezmodyfikowanej i ASL¹⁰⁰-A37 wskazuje na to, że
obecność CAA7 zapobiega tworzeniu zmutacji struktury pełni obserwowanej w przypadku
cząsteczki ASL¹⁰⁰-A.

Zadanie z mutacji danych poręgi 34 (H¹ i N¹H) dla cząsteczki ASL¹⁰⁰-A37
zmi kątowa struktura nie wykazywała wzrostu się typowego dla modyfikacji struktury RNA
modyfikacji U¹-test jest odpowiedniość na powstanie zmian kątów (-180°)

szkieletu fosfordiestrowego tRNA w miejscu sąsiadującym ze strony 5' z trzema nukleotydami antykodonu. Oddziaływania warstwowe rozciągają się wówczas od strony 3' trzonu ramienia antykodonowego do miejsca, gdzie ostry zwrot zmienia kierunek łańcucha oligorybonukleotydowego. Reszty fosforanowe w miejscu 33p34 mają wówczas charakterystyczne położenie w widmie ^{31}P NMR. W widmie wykonanym w wodzie nie zaobserwowałam żadnego sygnału, który świadczyłby o istnieniu, istotnego dla identyfikacji motywu *U-turn*, wiązania wodorowego pomiędzy iminowym protonem N3H urydyny U33 a grupą fosforanową reszty U36. Z uwagi na silne nakładanie się sygnałów rezonansowych nie mogłam potwierdzić także braku lub obecności innego charakterystycznego dla tego motywu sygnału korelacyjnego w widmie NOESY pomiędzy protonem anomerycznym H1' reszty U33 a protonem H6 reszty U35. Jednak brak typowego, silnie odsłanianego sygnału od atomu fosforu U33pU34 w widmie ^{31}P NMR dowodził absencji kluczowego dla struktury *U-turn* ostrego zwrotu kierunku szkieletu fosfordiestrowego. Wyniki te świadczą pośrednio o ważnej roli modyfikowanych nukleozydów w pozycji *wobble* i tłumaczą fakt strukturalnego niedopasowania reszty U34 (patrz powyżej) do pełnego zaangażowania w oddziaływanie antykodon-kodon.

Rozwój wydarzeń pokazał, że pośpiech prof. Agrisa w badaniach nad strukturą $\text{ASL}^{\text{Lys}3}$ -t⁶A37 był usprawiedliwiony. Trzy tygodnie przed ukazaniem się naszej pracy (V), została opublikowana, także w *Biochemistry*, struktura NMR ramienia antykodonowego tRNA^{Lys3} z *E.coli* zawierającego wszystkie modyfikacje (*mnm*⁵*s*²*U*, t⁶*A*, Ψ) (63). W tym samym czasie opublikowano również wyniki analizy krystalograficznej (rozdzielczość 3Å) struktury natywnego tRNA^{Lys3} z wątroby wołu i królika (64). W obu, w pełni modyfikowanych strukturach, zaobserwowano motyw *U-turn*. Cząsteczki tRNA^{Lys} posiadają nietypowy tryplet antykodonu UUU. W strukturze NMR cząsteczki $\text{ASL}^{\text{Lys}3}$ zawierającej nukleozyd *mnm*⁵*s*²*U* w pozycji 34 i t⁶A w pozycji 37 zaobserwowano, że motyw *U-turn* wpływa na usztywnienie konformacji modyfikowanej urydyny w pozycji *wobble* (U34). Przyczynia się to do stabilizacji słabych oddziaływań warstwowych pomiędzy resztami U35 i U36. Modyfikacje nukleozydów powodujące usztywnienie konformacji pętli są szczególnie istotne dla funkcjonowania tRNA^{Lys}, ponieważ oddziaływania kodon - antykodon są mediowane poprzez trzy relatywnie słabe oddziaływania par zasad zawierających urydynę. W strukturze krystalograficznej natywnej cząsteczki tRNA^{Lys3} grupa tionowa modyfikowanej urydyny U34 znajduje się dokładnie



nad aromatycznym pierścieniem reszty U35, co jak wiadomo z badań w kryształach (65) jest czynnikiem wzmacniającym oddziaływanie warstwowe.

Porównanie wyników badań otrzymanych przez mnie z wynikami pozostałych grup badawczych wskazuje, że: (i) reszta hipermodyfikowanego nukleozydu t^6A aktywnie moduluje oddziaływanie zasada-zasada zarówno o charakterze wiązań wodorowych jak i oddziaływań warstwowych, oraz (ii) hipermodyfikacja $A \rightarrow t^6A$ ($ASL^{Lys3} \rightarrow ASL^{Lys3}-t^6A$), a więc obecność jednej tylko modyfikacji t^6A37 nie jest wystarczająca dla uzyskania funkcjonalnie ważnej struktury pętli antykodonu. Dla stabilizacji oddziaływań wewnątrz pętli niezbędna jest również modyfikacja urydyny w pozycji 34: mcm^5s^2U lub mnm^5s^2U . Bogata w reszty uracylowe sekwencja prowadzi również prawdopodobnie do występowania tylko słabych oddziaływań warstwowych (65) w obrębie pętli cząsteczki $ASL^{Lys3}-t^6A37$.

Warto dodać, że reszta treonylowa wydaje się być naturalnym centrum kompleksowania jonów metali. Mimo, że wyniki z naszego laboratorium sugerują możliwość stabilizacji cząsteczki t^6A jonami Mg^{2+} (66), to wykazanie wagi tego odkrycia na poziomie $ASL^{Lys3}-t^6A$ było dotąd niemożliwe. Tym samym struktura krystaliczna $tRNA^{Phe}$ (67,68) pozostaje jedynym przykładem udowodnionego centrum chelatowania jonu Mg^{2+} w rejonie pętli antykodonowej.

5. UWAGI KOŃCOWE

Mimo ogromnego postępu w analizie strukturalnej kwasów nukleinowych w kryształach i w roztworze, nasza znajomość tych struktur jest nadal fragmentaryczna. Odzwierciedleniem tych trudności może być liczba struktur NMR kwasów nukleinowych zdeponowanych w bazie danych PDB (69). Obecnie znajduje się tam zaledwie około 100 struktur RNA i 200 DNA otrzymanych za pomocą spektroskopii NMR, na ogólną liczbę ponad 14000 wszystkich zdeponowanych struktur białek, peptydów, kwasów nukleinowych i kompleksów kwas nukleinowy-białko.

Pracownia Chemii Strukturalnej Kwasów Nukleinowych IChB PAN, w której pracuję, według mojej wiedzy jest jedyną w Polsce, która rozwija badania funkcjonalnie ważnych domen RNA metodami NMR. Jak dotychczas udział naszej Pracowni w liczbie zdeponowanych w bazie PDB struktur kwasów nukleinowych, jest skromny (70). Przyczyna tego leżała w braku bezpośredniego dostępu do bardzo kosztownych przyrządów, jakimi są odpowiednio wyposażone i przystosowane do badań strukturalnych biomolekuł spektrometry NMR. Aspekt ten jest szczególnie istotny w przypadku badań łatwo degradujących się preparatów RNA. Z tych względów wszystkie pomiary fragmentów RNA wykonałam za granicą na spektrometrach o częstotliwościach dla protonów 500-600 MHz. Obecnie sytuacja ta ulega znacznej poprawie, gdyż przybyło w Polsce, również w Poznaniu, wysokorozdzielczych spektrometrów NMR.

Należy w tym miejscu podkreślić, że o ile ośrodek poznański nie dysponował odpowiednimi przyrządami NMR, to od kilku lat miałam dostęp do mocy obliczeniowych i specjalistycznego oprogramowania (pakiety Biosym/MSI, SYBYL) w afiliowanym przy IChB PAN Poznańskim Centrum Superkomputerowo-Sieciowym. Pozwoliło mi to na kontynuację w kraju, realizowanych za granicą projektów badawczych (III,IV,V).

6. CYTOWANA LITERATURA

1. Pardi A.
Methods Enzymol., **261**,350-380, 1995
2. Prestegard J.H.
Nat. Struct. Biol., **5** Suppl, 517-522, 1998
3. Pervushin K., Riek R., Wider G., Wuthrich K.
Proc. Natl. Acad. Sci. USA., **94**, 12366-12371, 1997
4. Pervushin K., Ono A., Fernandez C., Szyperski T., Kainosho M., Wuthrich K.
Proc. Natl. Acad. Sci. USA., **95**, 14147-14151, 1998
5. Batey R.T., Battiste J.L., Williamson J.R.
Methods Enzymol., **261**, 300-322, 1995
6. Kainosho M.
Nat. Struct. Biol., **4** Suppl, 858-861, 1997
7. Ono A., Tate S. Ishido Y., Kainosho M.,
J. Biomol. NMR, **4**, 581-586, 1994
8. Milecki J., Zamaratski E., Maltseva T.V., Foldesi A., Adamiak R.W.,
Chattopadhyaya J.
Tetrahedron, **55**, 6603-6622, 1999
9. Breen A.P., Murphy J.A.
Free Rad. Biol. Med., **18**, 1033-77, 1995
10. Modrich P.
Annu. Rev. Biochem., **56**,435-66, 1987
11. Modrich P.
Basic Life Science., **38**, 303-10, 1986
12. Radman M., Wagner R.
Annu. Rev. Genetics, **20**,523-38, 1986
13. Mellac S., Fazakerley G.V., Sowers L.C.
Biochemistry, **32**, 7779-7786, 1993
14. Anand NN., Brown DM., Salisbury SA.
Nucleic Acids Res., **15**, 8167-76, 1987
15. Budowski , E.I.
Progr. Nucleic Acids Res. Mol. Biol., **16**, 125-188, 1976

6 CYTOSOL LITIGATION

1. Parfi A.
Mol Biol Cell, 5:150-152, 1992
2. Prasad S.H.
Mol Biol Cell, 5:17-22, 1994
3. Prasad S.H., Rief R., Winer D., Wuttsch K.
Proc Natl Acad Sci USA, 91:1230-1233, 1994
4. Prasad S.H., Rao A., Prasad S.H., Wuttsch K., Wuttsch K.
Proc Natl Acad Sci USA, 91:1415-1418, 1994
5. Prasad S.H., Wuttsch K., Wuttsch K.
Mol Biol Cell, 5:300-322, 1994
6. Prasad S.H.
Mol Biol Cell, 4:Suppl, 822-821, 1993
7. Prasad S.H., Wuttsch K., Wuttsch K.
J Biol Chem, 269:1521-1524, 1994
8. Prasad S.H., Wuttsch K., Wuttsch K., Wuttsch K., Wuttsch K.
Chromosoma, 55:603-612, 1995
9. Prasad S.H., Wuttsch K.
Proc Natl Acad Sci USA, 91:1033-1037, 1994
10. Prasad S.H.
Ann Rev Biochem, 56:415-44, 1987
11. Prasad S.H.
Basic Life Sci, 38:103-10, 1986
12. Prasad S.H., Wuttsch K.
Ann Rev Biochem, 56:23-38, 1987
13. Prasad S.H., Wuttsch K., Wuttsch K.
Biochem J, 215:779-782, 1993
14. Prasad S.H., Wuttsch K., Wuttsch K.
Mol Biol Cell, 4:1515-1518, 1993
15. Prasad S.H.
Proc Natl Acad Sci USA, 91:122-124, 1994

16. Shugar D., Huber C.P., Birnbaum G.I.
Biochim. Biophys. Acta., **447**, 274-84, 1976
17. Morozov, Y.V., Savin F.A. Chechov V.O. Budovskii E.I. Yakovlev D.Y.
J. Photochem., **20**, 229-252, 1982
18. Van Meervelt L., Moore M.H., Lin P.K., Brown D.M., Kennard O.
J. Mol. Biol., **216**, 773-781, 1990
19. Brown D.M., Hewlins M.J., Schell P.
J. Chem. Soc.-Perkin Trans. 1., **15**, 1925-1929, 1968
20. Kierdaszuk B., Shugar D.
Biophys. Chem., **17**, 285-295, 1983
21. Kierdaszuk B., Stolarski R., Shugar D.
Eur. J. Biochem., **130**, 559-564, 1983
22. Lin P.K., Brown D.M.
Nucleic Acids Res., **17**, 10373-10383, 1989
23. Stone M.J., Nedderman A.N., Williams D.H., Lin P.K., Brown D.M.
J. Mol. Biol., **222**, 711-723, 1991
24. Nedderman A.N., Stone M.J., Williams D.H., Lin P.K., Brown D.M.
J. Mol. Biol., **230**, 1068-1076, 1993
25. Hikima T., Hossain M.T., Chatake T., Ueno Y., Matsuda A., Takenaka A.
Nucleic Acids Res. Symp. Series, **42**, 49-50, 1999
26. Theil EC.
Biochem. J., **304**, 1-11, 1994
27. Sierzputowska-Gracz H., McKenzie RA., Theil EC.
Nucleic Acids Res., **23**, 146-153, 1995
28. Laing L.G., Hall K.B.
Biochemistry, **35**, 13586-13596, 1996
29. Zuker M.
Computer Analysis of Sequence Data
A.M. Griffin and H.G. Griffin eds.
Methods in Molecular Biology, Humana Press Inc., 267-294, 1994
30. Puglisi J.D., Tan R., Calnan B.J., Frankel A.D., Williamson J.R.
Science, **257**, 76-80, 1992
31. Battiste J.L., Tan R., Frankel A.D., Williamson J.R.
Biochemistry, **33**, 2741-2747, 1994

16. Shugar D, Huber C.P., *Biochim Biophys Acta*, 467, 245-252 (1976)

17. Morozov, Y.V., Savin F.A. *J. Mol. Biol.*, 20, 229-252 (1977)

18. Van Meerwijk L., Moore M.H. *J. Mol. Biol.*, 216, 773-781 (1990)

19. Brown D.M., Hewlin M.L. *J. Chem Soc Chem Commun*, 1987, 1257-1258

20. Kierbaszuk B., Shugar D. *Biochim Biophys Acta*, 17, 285-290 (1976)

21. Kierbaszuk B., Stolarski R., Szymanski J. *Acta Biochim*, 138, 229-236 (1987)

22. Lin P.K., Brown DM. *Nucleic Acids Res.*, 17, 1037-1042 (1989)

23. Stone M.L., Nederman A.M. *J. Mol. Biol.*, 222, 711-723 (1991)

24. Nederman A.M., Stone M.L. *J. Mol. Biol.*, 230, 1052-1072 (1992)

25. Hittina T., Hostain M.T., *Nucleic Acids Res Suppl*, 2000, 100

26. Trail HC. *Biochem J*, 304, 1-11 (1994)

27. Sierpniowska-Guzek H., *Nucleic Acids Res*, 23, 146-152 (1995)

28. Iking L.O., Hall K.B. *Biochemistry*, 35, 15286-15291 (1996)

29. Zickler M. *Computer Analysis of Sequences*, A.M. Griffin and H.O. Griffith, *Methods in Molecular Biology*, 1997

30. Foghail J.D., Tan K., *Science*, 257, 76-80 (1992)

31. Buttrose J.L., Tan R., *Biochemistry*, 33, 2741-2745 (1994)

32. Wimberly B., Varani G., Tinoco I.Jr.
Biochemistry, **32**, 1078-87, 1993
33. Szewczak A.A., Moore P.B., Chang Y.L., Wool I.G.
Proc. Natl. Acad. Sci. USA., **90**, 9581-9585, 1993
34. Zuker M., Mathews D.H., Turner D.H.
In *RNA Biochemistry and Biotechnology*, 11-43, J. Barciszewski & B.F.C. Clark, eds., NATO ASI Series, Kluwer Academic Publishers, 1999
35. Mathews D.H., Sabina J., Zuker M., Turner D.H.
J. Mol. Biol., **288**, 911-940, 1999
36. Major, F., Turcotte, M., Lapalme, G., Fillion, E., and Cedergren, R.
Science, **253**, 1255-1260, 1991
37. Gautheret D., Major F., Cedergren R.
J. Mol. Biol. **229**, 1049-64, 1993
38. Major, F., Gautheret, D., and Cedergren, R.,
Proc. Natl. Acad. Sci. USA, **90**, 9408-9412, 1993
39. Leclerc F., Cedergren R., Ellington AD.
Nat. Struct. Biol., **1**, 293-300, 1994
40. Address K.J., Basilion J.P., Klausner R.D., Rouault T.A., Pardi, A.
J. Mol. Biol., **274**, 72-83, 1997
41. Limmer S.
Prog. Nucl. Acid Res. Mol. Biol. **57**, 1-39, 1997
42. Pley H.W., Flaherty K.M., McKay D.B.
Nature, **372**, 68-74, 1994
43. Cate J.H., Doudna J.A.
Structure, **4**, 1221-1229, 1996
44. Correll C.C., Freeborn B., Moore P.B., Steitz T.A.
Cell, **91**, 705-12, 1997
45. Feig A.L., Uhlenbeck O.C.
The RNA World, Second Edition
Edited by Gesteland R.F., Cech T.R., Atkins J. F.
Cold Spring Harbor Laboratory Press, 1999
46. Hermann T., Westhof E.
Structure, **6**, 1303-14, 1998

32	Wang, H. C. (1978)	100
33	Wang, H. C. (1979)	100
34	Wang, H. C. (1980)	100
35	Wang, H. C. (1981)	100
36	Wang, H. C. (1982)	100
37	Wang, H. C. (1983)	100
38	Wang, H. C. (1984)	100
39	Wang, H. C. (1985)	100
40	Wang, H. C. (1986)	100
41	Wang, H. C. (1987)	100
42	Wang, H. C. (1988)	100
43	Wang, H. C. (1989)	100
44	Wang, H. C. (1990)	100
45	Wang, H. C. (1991)	100
46	Wang, H. C. (1992)	100

47. Cowan J.A.
J. Inorg. Biochem., **49**, 171-175, 1993
48. Kieft J.S., Tinoco I. Jr.
Structure, **5**, 713-21, 1997
49. Kim S.H., Suddath F.L., Quigley G.J., McPherson A., Sneden D., Kim J.J., Kim SH., Rich A.
Nature, **248**, 20-24, 1974
50. Robertus J.D., Ladner J.E., Finch J.T., Rhodes D., Brown R.S., Clark B.F., Klug A.
Nature, **250**, 546-551, 1974
51. Adamiak R.W., Biala E., Grzeskowiak K., Kierzek R., Kraszewski A., Markiewicz W.T., Okupniak J., Stawinski J., Wiewiorowski M.
Nucleic Acids Res., **5**, 1889-1905, 1978
52. Boudou V., Langridge J., Van Aerschot A., Hendrix C, Millar A., Weiss P., Herdewijn P.
Helv. Chim. Acta., **83**, 152-161, 2000
53. Sundaram M., Crain P.F., Davis D.R.
J. Org. Chem., **65**, 5609-5614, 2000
54. Grosjean M.
Modification and Editing of RNA
Editors: Henri Grosjean, and Rob Benne, 1998 ASM Press
55. Ashraf S.S., Ansari G., Guenther R., Sochacka E., Małkiewicz A., Agris PF.
RNA-Publ. RNA Soc., **5**, 503-511, 1999
56. Yarian C., Marszałek M., Sochacka E., Małkiewicz A., Guenther R., Miskiewicz A. Agris PF.
Biochemistry, **39**, 13390-13395, 2000
57. Stejskal E.O., Tanner, J.E.
J. Chem. Phys., **42**, 288-292, 1965
58. Lapham J., Rife J., Moore P.B., Crothers D.M.
J. Biomol. NMR, **10**, 255-262, 1997
59. Butcher S.E., Dieckmann T., Feigon J.
J. Mol. Biol., **268**, 348-358, 1997
60. Durant P.C. Davis D.R.
J. Mol. Biol., **285**, 115-31, 1999
61. Legault P., Pardi A.
J. Am. Chem. Soc., **116**, 8390-8391, 1994

47. Coan J.A. *J. Inorg. Biochem.*, **42**, 171-172, 1992

48. Kish J.S., Thero J. Jr. *Structure*, **2**, 713-71, 1997

49. Kim B.H., Boubath F.L., Gulguy O.J., Moberg A., Sunden D., Kim U., Kim S.H., Rish A. *Nature*, **348**, 39-44, 1994

50. Roberts J.D., Fisher J.E., Fisher I.T., Rhodes D., Brown R.E., Clark B.P., King A. *Nature*, **330**, 544-551, 1974

51. Adams R.W., Hsiao E., Grzeskowiak K., Kierulff R., Krausz A., Malinowski W.T., Chapman L., Sliemers L., Wozniowski M. *Acta Cryst. Ser. B*, **18**, 1884-1902, 1992

52. Bostov V., Langridge J., Von Ammon A., Hendrix C., Miller A., Weiss P., Hertzberg P. *Acta Cryst. Ser. B*, **28**, 152-161, 2000

53. Sunden M., Crain P.R., Davis D.R. *J. Org. Chem.*, **65**, 5600-5614, 2000

54. Grogan M. *Medicines and Editing of RNA*. Editors: Hans Grogan, and Rob Howe, 1998. ASM Press

55. Albert S.F., Averbach G., Gendry R., Seebach E., Malinowski A., Agre P. *EMBO J.*, **16**, 507-511, 1997

56. Yama C., Malinowski M., Seebach E., Malinowski A., Gendry R., Malinowski A. *Biochemistry*, **39**, 13300-13305, 2000

57. Seibel E.O., Turner J.B. *J. Chem. Phys.*, **42**, 588-592, 1965

58. Lipton J., Rife J., Moore F.B., Grollman D.M. *J. Blood. Med.*, **10**, 232-242, 1997

59. Butler S.E., Blockman T., Feigon J. *J. Mol. Biol.*, **302**, 344-354, 1997

60. Durrant P.C., Davis D.R. *J. Mol. Biol.*, **282**, 115-121, 1999

61. Legault P., Furdik A. *J. Am. Chem. Soc.*, **118**, 6300-6307, 1994

62. Legault P., Pardi A.
J. Am. Chem. Soc., **119**, 6621-6628, 1997
63. Sundaram M., Durant P.C., Davis D.R.
Biochemistry, **39**, 12575-12584, 2000
64. Benas P., Bec G., Keith G., Marquet R., Ehresmann C., Ehresmann B., Dumas P.
RNA-Publ. RNA Soc., **6**, 1347-1355, 2000
65. Saenger W.,
Principles of Nucleic Acid Structure
Springer Verlag New York 1984.
66. Varnagy K., Jeżowska-Bojczuk M., Swiatek J., Kozłowski H., Sovago I.,
Adamiak R.W.
J. Inorg. Biochem., **40**, 357-63, 1990
67. Holbrook S.R., Sussman J.L., Warrant R.W., Church G.M., Kim S.H.
Nucleic Acids Res., **4**, 2811-20, 1977
68. Jack A., Ladner J.E., Rhodes D., Brown R.S., Klug A.
J. Mol. Biol., **111**, 315-28, 1977
69. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H.,
Shindyalov I.N., Bourne P.E.
Nucleic Acids Res., **28**, 235-242, 2000
70. Popenda M., Biała E., Milecki J., Adamiak R.W.
Nucleic Acids Res., **25**, 4589-4598, 1997

- 1. The first part of the book is devoted to a general introduction to the subject of the history of the world.
- 2. The second part of the book is devoted to a detailed account of the history of the world from the beginning of time to the present day.
- 3. The third part of the book is devoted to a detailed account of the history of the world from the beginning of time to the present day.
- 4. The fourth part of the book is devoted to a detailed account of the history of the world from the beginning of time to the present day.
- 5. The fifth part of the book is devoted to a detailed account of the history of the world from the beginning of time to the present day.
- 6. The sixth part of the book is devoted to a detailed account of the history of the world from the beginning of time to the present day.
- 7. The seventh part of the book is devoted to a detailed account of the history of the world from the beginning of time to the present day.
- 8. The eighth part of the book is devoted to a detailed account of the history of the world from the beginning of time to the present day.
- 9. The ninth part of the book is devoted to a detailed account of the history of the world from the beginning of time to the present day.
- 10. The tenth part of the book is devoted to a detailed account of the history of the world from the beginning of time to the present day.

I.

.I

Base-pair Induced Shifts in the Tautomeric Equilibrium of a Modified DNA Base

G. Victor Fazakerley†

*Service de Biochimie et de Génétique Moléculaire, Bât. 142
Département de Biologie Cellulaire et Moléculaire, Centre d'Etudes de Sacaly
91191 Gif-sur-Yvette Cedex, France*

Zofia Gdaniec

*Service de Biochimie et de Génétique Moléculaire, Bât. 142
Département de Biologie Cellulaire et Moléculaire, Centre d'Etudes de Sacaly
91191 Gif-sur-Yvette Cedex, France
and Institute of Bioorganic Chemistry, Polish Academy of Sciences
Noskowskiego 12/14, 61-704 Poznan, Poland*

and Lawrence C. Sowers†

*Division of Pediatrics, City of Hope National Medical Center
Duarte, CA 91010, U.S.A.*

(Received 18 June 1992; accepted 4 December 1992)

We have examined the base-pairing properties of N⁴-methoxycytosine (mo⁴C), a mutagenic base analog, in DNA by nuclear magnetic resonance spectroscopy. Unlike standard bases, the tautomeric equilibrium of mo⁴C could be strongly influenced by base-pair formation. Paired with A, mo⁴C is found predominantly in the imino configuration in Watson-Crick geometry. However, when paired with G, two structurally distinct configurations are observed in equilibrium with one another. In one configuration, mo⁴C is in the amino form paired with G in Watson-Crick geometry. In the second species, mo⁴C is in the imino configuration paired with G in a wobble geometry. This is the first demonstration of base-pair induced tautomeric shifts in DNA and supports the hypothesis that rare tautomeric forms may be involved in mutagenesis.

Keywords: mutagenesis; n.m.r.; DNA; tautomerism; oligonucleotide

The three-dimensional structure of DNA proposed by Watson & Crick (1953) provided a molecular explanation for the transfer of genetic information in biological systems. In the same paper, it was suggested how errors in information transfer *via* spontaneous genetic mutations could occur. It was proposed that base-mispairing would result if a base was present in its less likely tautomeric form, one in which a proton is transferred from one site to another on the base. Rare tautomer forms are ubiquitously present, but in very low concentration. The experimental approach to search

for these forms has been to construct oligonucleotide duplexes containing mispairs and to determine whether these mispairs exist in the predicted rare tautomeric configurations. To date, experimental studies on such systems by either X-ray crystallography or nuclear magnetic resonance (n.m.r.†) have identified both wobble (Hare *et al.*, 1986; Quignard *et al.*, 1987; Sowers *et al.*, 1988, 1989; Carbonnaux *et al.*, 1990) and ionized (Patel *et al.*, 1984; Sowers *et al.*, 1986, 1988, 1989; Wang *et al.*,

† Abbreviations used: n.m.r., nuclear magnetic resonance; mo⁴C, N⁴-methoxycytosine; p.p.m., parts per million; NOESY, 2-dimensional NOE spectroscopy; NOE, nuclear Overhauser effect.

† Authors to whom all correspondence should be addressed.

Base-pair Induced

Department

Department

and

Division

We have examined the base pairing in the DNA double helix. The interaction between the two strands is primarily due to hydrogen bonding between the nitrogenous bases. The major groove and minor groove are formed by the arrangement of the bases. The major groove is wider and deeper than the minor groove. The major groove is the site for the binding of many proteins and drugs. The minor groove is the site for the binding of some proteins and drugs. The major groove is also the site for the binding of the major groove binding proteins. The minor groove is also the site for the binding of the minor groove binding proteins. The major groove is also the site for the binding of the major groove binding proteins. The minor groove is also the site for the binding of the minor groove binding proteins.

The three-dimensional structure of DNA is a double helix. The two strands are held together by hydrogen bonds between the nitrogenous bases. The major groove and minor groove are formed by the arrangement of the bases. The major groove is wider and deeper than the minor groove. The major groove is the site for the binding of many proteins and drugs. The minor groove is the site for the binding of some proteins and drugs. The major groove is also the site for the binding of the major groove binding proteins. The minor groove is also the site for the binding of the minor groove binding proteins. The major groove is also the site for the binding of the major groove binding proteins. The minor groove is also the site for the binding of the minor groove binding proteins.

J. Authors to whom all correspondence should be addressed.

1991) base-pairs, however, evidence for rare tautomers has not been presented in any of these studies. Here, we present data which show that a configurational equilibrium *via* tautomerization can occur within a base-pair and that tautomeric equilibrium can be strongly influenced by base-pair formation.

The experimental challenge to identify rare tautomeric forms derives from the fact that, in most cases, one tautomeric form is highly preferred. The rare forms only exist at the level of 1 in 10^4 to 10^5 for the standard bases (Topal & Fresco, 1976; Fresco *et al.*, 1980). The presence of tautomeric forms at this level may be biologically significant at their low concentration and short lifetime put them outside the detection limit of most physical methods. In some cases, however, it is known that base modifications can significantly affect the tautomeric equilibrium for an isolated base. Substitution of one hydrogen atom of a cytosine amino group by an electron withdrawing substituent such as amino, hydroxy or methoxy groups can induce formation of the imino form with tautomeric equilibrium constants orders of magnitude lower than those for the standard DNA bases (Brown *et al.*, 1968; Spengler & Singer, 1981; Kierdaszuk *et al.*, 1983; Anand *et al.*, 1987; Stone *et al.*, 1991). One of these, 4-methoxycytosine (mo^4C), is the product of the reaction of methoxyamine, a known mutagen, with cytosine. It has been shown that mo^4C pairs predominantly like thymine in DNA transcription experiments (Spengler & Singer, 1981).

A recent n.m.r. study (Stone *et al.*, 1991) has shown that mo^4C , when paired with A, is found in the imino conformation with Watson-Crick geometry, consistent with previous biochemical studies. Due to the low tautomerization energy, it had been expected that mo^4C would pair with G in the amino conformation. However, an X-ray study of a Z-form duplex containing an $mo^4C \cdot G$ pair found only the imino form with the G in wobble geometry (van Meervelt *et al.*, 1990).

We have continued to study this system as it is that which is most likely to exhibit base-pair induced tautomeric shifts. Our studies were carried out in solution so that there are no constraints from the crystal lattice. The two duplexes studied are shown in Figure 1 and these oligonucleotides were prepared by standard phosphoramidite methods as

G ₁	A ₂	G ₃	mo^4C	G ₅	G ₆	C ₇	Duplex $mo^4C \cdot A$
G ₁₄	T ₁₃	C ₁₂	A ₁₁	C ₁₀	C ₉	G ₈	
C ₁	A ₂	G ₃	mo^4C	G ₅	G ₆	C ₇	Duplex $mo^4C \cdot G$
G ₁₄	T ₁₃	C ₁₂	G ₁₁	C ₁₀	C ₉	G ₈	

Figure 1. Sequences of the 2 duplexes studied. The n.m.r. samples were prepared by annealing equimolar amounts of the appropriate pairs of strands in 50 mM-NaCl, 10 mM-sodium phosphate buffer (pH 7.0) dissolved in 90% H_2O , 10% 2H_2O at 80°C followed by slow cooling. Spectra were recorded at 600 MHz on a Bruker AMX 600 spectrometer.

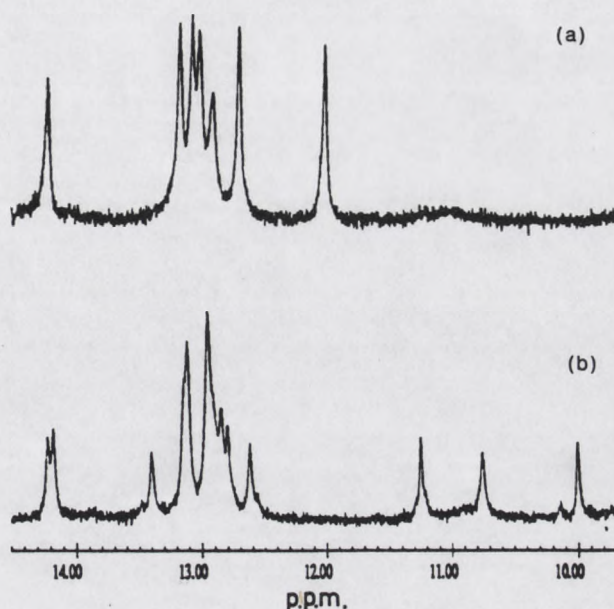


Figure 2. Spectra of the 2 duplexes at 1°C showing the region of the Watson-Crick base-pair imino protons and also the amino or imino proton of mo^4C . (a) duplex $mo^4C \cdot A$ and (b) $mo^4C \cdot G$. The imino proton of the mo^4C base-paired with A is observed at 12.05 p.p.m.

described previously (Anand *et al.*, 1987). The low field 1H spectra, corresponding to exchangeable protons of each duplex are shown in Figure 2(a) and (b), respectively. Resonance assignments were made by standard methods (Hare *et al.*, 1983; Frechet *et al.*, 1983; Feigon *et al.*, 1983; Fazakerley *et al.*, 1984; Scheek *et al.*, 1984). For the $mo^4C \cdot A$ duplex, only the imino configuration of mo^4C is observed with the imino proton resonating at 12.05 parts per million (p.p.m.) (Fig. 2(a)). This chemical shift indicates that the $mo^4C \cdot A$ base-pair has Watson-Crick geometry. The broad hump observed at approximately 11 p.p.m. probably arises from a slight excess of one strand but it can not be excluded that this corresponds to the amino proton of the other tautomer of mo^4C , the two conformations being in slow exchange on the n.m.r. time-scale. Although we have observed some differences in this system with those reported recently (Stone *et al.*, 1991), our conclusions are essentially the same.

The spectrum of the $mo^4C \cdot G$ duplex is visibly more complicated with variable peak integrals, indicating the presence of more than one species (Fig. 2(b)). Peak assignment has been carried out by two-dimensional nuclear Overhauser effect spectroscopy (NOESY). The two high field resonances, at 10.00 and 10.76 p.p.m., show strong interactions between each other (Fig. 3) indicating close spatial proximity. However, close intrabase contacts would be expected in both the imino and amino conformations (Fig. 4(a) and (b), respectively). The two configurations can be distinguished because in the imino form, the G11 amino group is not hydrogen-bonded, as in wobble base-pairs, and is free to

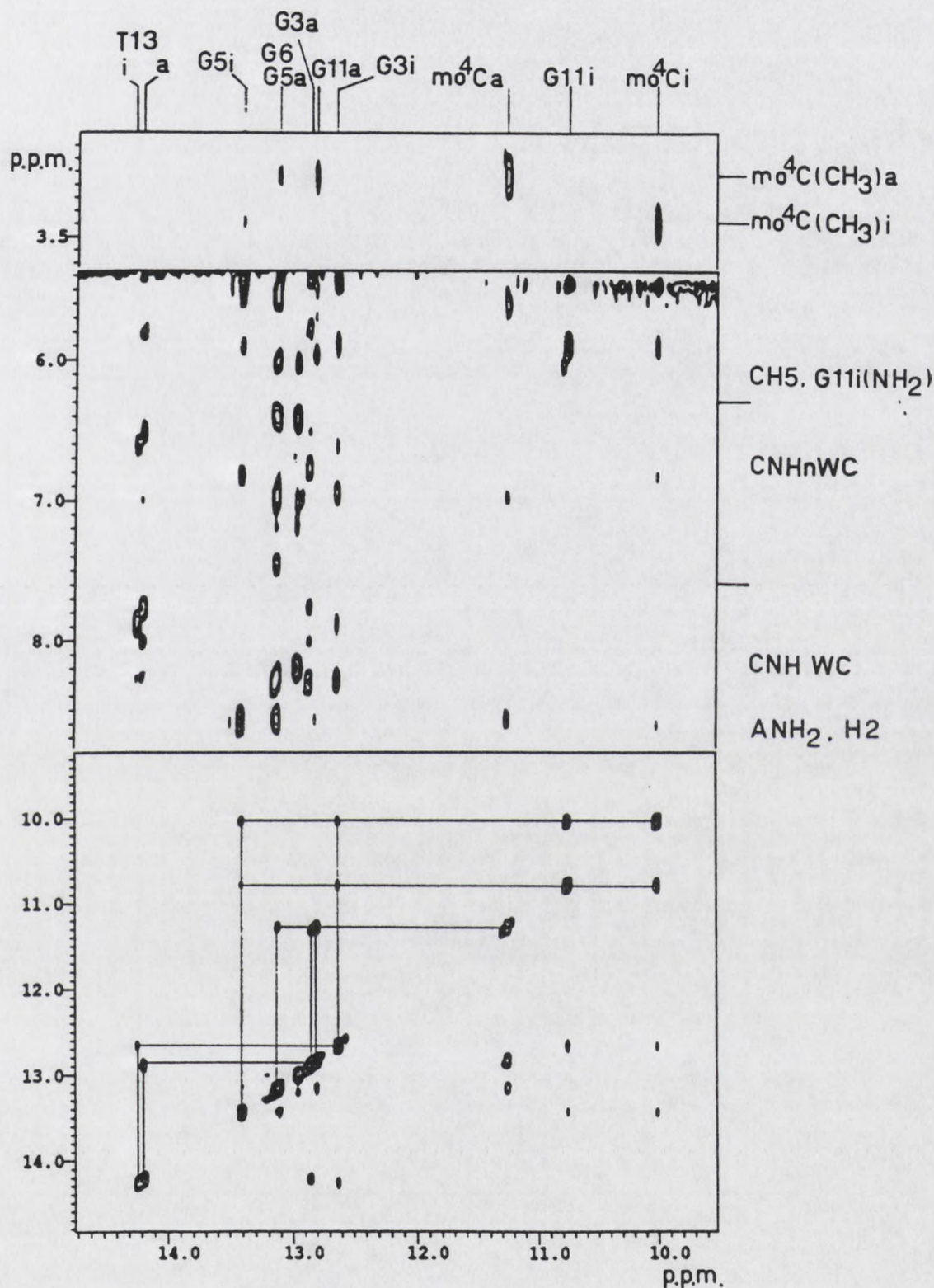


Figure 3. Three regions of the NOESY spectrum of $mo^4C \cdot G$. The assignment of the imino and mo^4C amino or imino protons is shown on the horizontal axis. Where separate resonances are observed for the 2 species these are labeled as follows: i, for where the mo^4C base is in the imino form and a, where this base is in the amino form. The lower part shows connectivities between the imino protons, and also the amino proton of mo^4C . The middle section shows interactions between these protons and the aromatic and C hydrogen-bonded (WC) and non-hydrogen bonded protons (nWC). The upper part of the Figure shows interactions with the mo^4C methyl resonance which has a different chemical shift for the 2 forms. Chemical shift regions for different types of protons are indicated on the right hand side of the Figure.

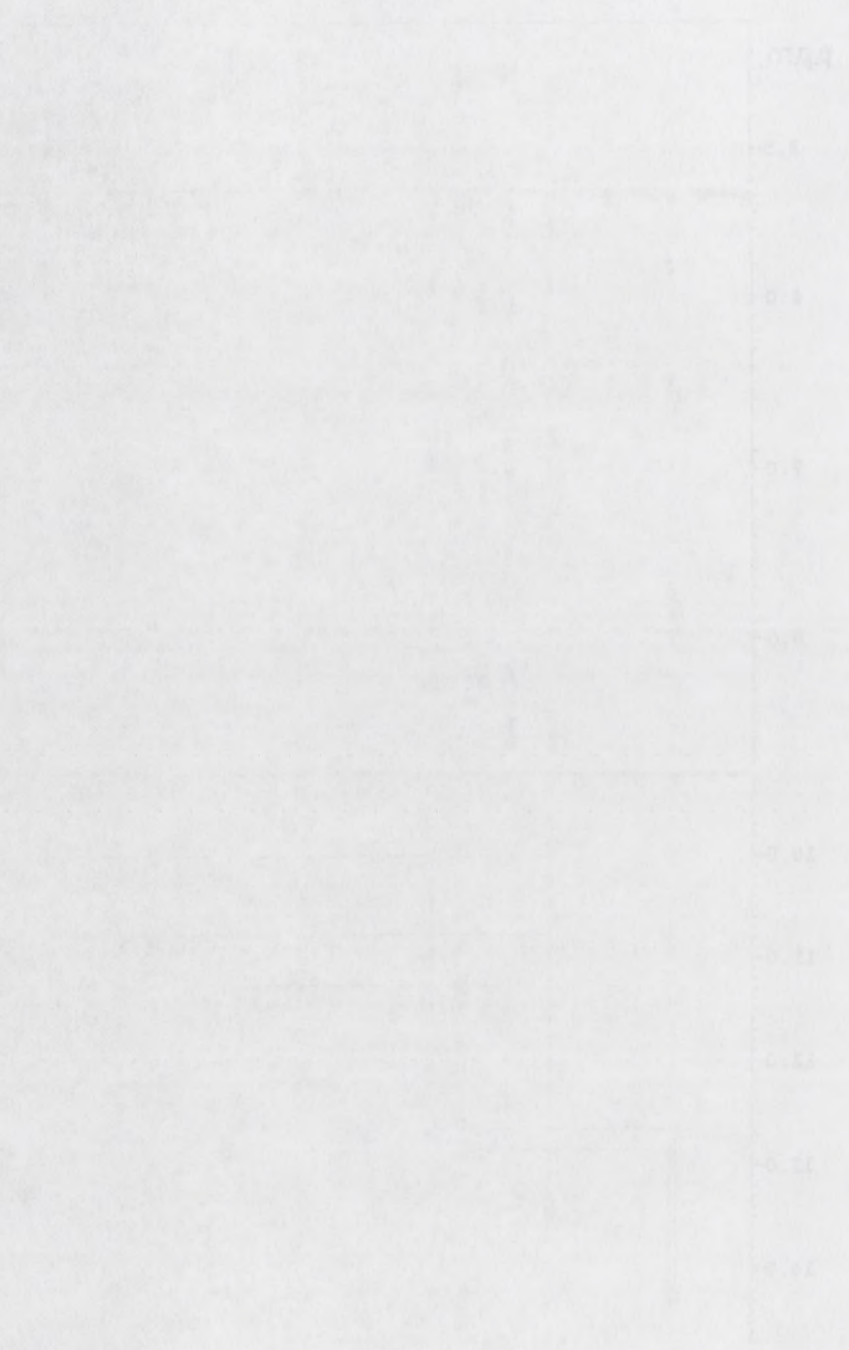


Figure 1. The
 graph shows the
 relationship between
 the variables
 plotted on the
 axes. The data
 points are
 scattered across
 the plot area.

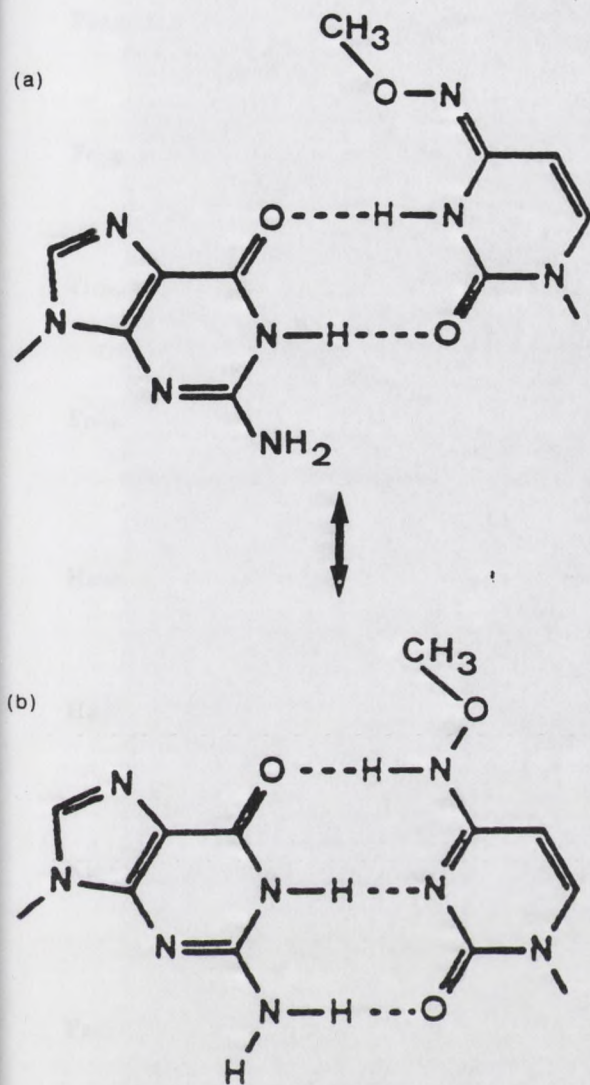


Figure 4. (a) The structure of the mo^4C -G base-pair in wobble geometry with the mo^4C base in the imino form, and (b) in Watson-Crick geometry with the mo^4C base in the amino form.

rotate, giving rise to a single resonance for the amino group at 5.8 to 6.0 p.p.m. (Carbonnaux *et al.*, 1990). On the other hand, for the amino form of mo^4C , the amino group of G11 will be hydrogen-bonded to the keto group of mo^4C . For G-C base-pairs, it has been shown (Fazakerley *et al.*, 1984) that at low temperature rotation of the amino group and interchange of the amino protons is in the intermediate exchange rate on the proton n.m.r. time-scale and these protons are generally not observable. This enables us to distinguish the two forms for mo^4C . In Figure 3 the resonances at 10.00 and 10.76 p.p.m. show nuclear Overhauser effects (NOEs) with a resonance at 5.90 p.p.m., which is assigned to the amino group of G11 in the imino form of the mo^4C -G base-pair. Further, the NOE from the resonance at 10.76 p.p.m. to the amino group of G11 is much stronger than that from the

resonance at 10.00 p.p.m. and thus the former can be assigned to the G11 imino proton. This relative assignment is confirmed in the upper part of Figure 3. The resonance at 10.00 p.p.m. shows a strong NOE with the $mo^4C(CH_3)$ group of this species whereas the resonance at 10.76 p.p.m. shows no NOE. This also indicates that the mo^4C methyl group spends at least part of its time *cis* to N-3 in the imino tautomer. In the lower part of Figure 3 we also observe a strong cross peak between resonances at 11.25 and 12.80 p.p.m. (although there is partial overlap with an interbase-pair NOE) and these are assigned to the amino proton of mo^4C and the imino proton of G11 in the amino tautomer. Again the two can be distinguished by the interaction with the $mo^4C(CH_3)$. The resonance at 11.25 p.p.m. shows the strongest interaction and can be assigned to the amino proton of mo^4C . All of these protons exhibit normal interactions with the imino protons of the G-C pairs on either side. At 20°C, we observe exchange cross peaks for the H-5 and H-6 protons of mo^4C in the amino and imino configurations (Fig. 4) in NOESY spectra (not shown). The ratio of the amino and imino configurations changes with temperature.

The data demonstrate that different tautomeric forms and base-pair configurations can exist simultaneously within a DNA base-pair and that tautomeric equilibrium can be strongly influenced by the paired base in the opposing DNA strand, consistent with the original proposal of Watson and Crick. For standard bases, the difference in the stability of the tautomers is so large that the rare tautomers are not observable. As the energy of ionization is lower than that of tautomerization at physiological pH, ionized base-pairs may preclude observation of the rare tautomeric forms in most cases (Sowers *et al.*, 1986, 1987).

The consensus of physical studies to date leads to the conclusion that most aberrant base-pairs are likely to exist as composites of multiple structures in solution. These structures include both ionized and unfavored tautomeric forms of the bases in either Watson-Crick or wobble configurations. The relative contribution of each component structure will be influenced by temperature, solvent pH and by the base sequence. The data presented here show that mechanisms involving rare tautomers should not be discounted in biological systems.

This work was supported in part by the National Institutes of Health (grants GM41336 and CA33572).

References

- Anand, N. N., Brown, D. M. & Salisbury, S. A. (1987). The stability of oligodeoxyribonucleotide duplexes containing degenerate bases. *Nucl. Acids Res.* **15**, 8167-8176.
- Brown, D. M., Hewlins, M. J. E. & Schell, P. (1968). The tautomeric state of N(4)-hydroxy and N(4)-aminocytosine derivatives. *J. Chem. Soc. (C)*, 1925-1929.
- Carbonnaux, C., Fazakerley, G. V. & Sowers, L. C. (1990). An NMR structural study of deaminated base-pairs in DNA. *Nucl. Acids Res.* **18**, 4075-4081.

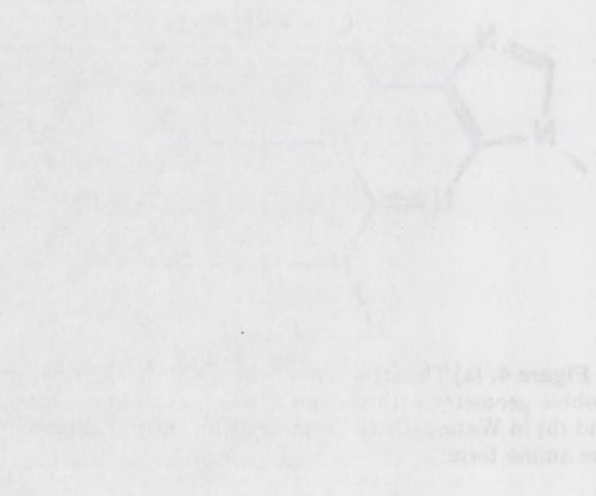
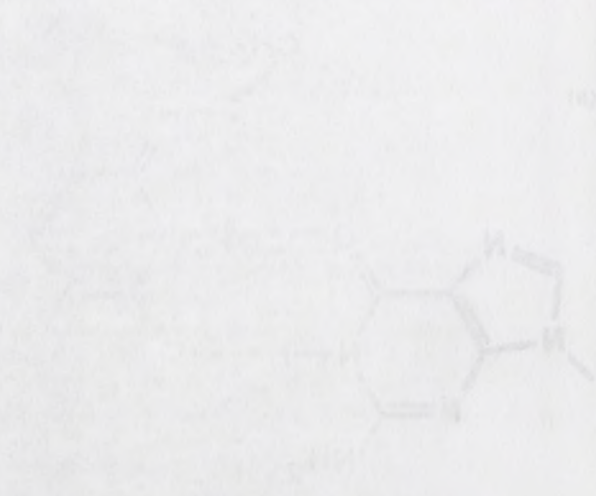


Figure 4. (a) The structure of the...
 (b) The structure of the...
 The same form...

The text in this section is extremely faint and largely illegible. It appears to be a caption or a short description of the chemical structures shown in the adjacent blocks. The text is oriented vertically on the page.

- Fazakerley, G. V., van der Marel, G. A., van Boom, J. H. & Guschlbauer, W. (1984). Helix opening in deoxyribonucleic acid from proton nuclear magnetic resonance study of imino and amino protons in d(CG)₃. *Nucl. Acids Res.* **12**, 8269-8279.
- Feigon, J., Leupin, W. A., Denny, W. & Kearns, D. R. (1983). Two-dimensional proton nuclear magnetic resonance investigation of the synthetic deoxyribonucleic acid decamer d(ATATCGATAT)₂. *Biochemistry*, **22**, 5943-5951.
- Frechet, D., Cheng, D. M., Kan, L.-S. & Tso, P. O. P. (1983). Nuclear Overhauser effect as a tool for the complete assignment of nonexchangeable proton resonances in short deoxyribonucleic acid helices. *Biochemistry*, **22**, 5194-5200.
- Fresco, J. R., Broitman, S. & Lane, A.-E. (1980). Base-mispairing and nearest-neighbor effects in transition mutations. In *Mechanistic Studies of DNA Replication and Genetic Recombination* (Alberts, B. & Fox, C. F., eds), pp. 753-768. Academic Press, New York.
- Hare, D., Wemmer, D. E., Chou, S.-H., Drobny, G. & Reid, B. R. (1983). Assignment of the non-exchangeable proton resonances of d(C-G-C-G-A-A-T-T-C-G-C-G) using two-dimensional nuclear magnetic resonance methods. *J. Mol. Biol.* **171**, 319-336.
- Hare, D., Shapiro, L. & Patel, D. (1986). Wobble dG·dT pairing in right-handed DNA: solution conformation of the d(C-G-T-G-A-A-T-T-C-G-C-G) duplex deduced from distance geometry analysis of nuclear Overhauser effect spectra. *Biochemistry*, **25**, 7445-7456.
- Kierdaszuk, B., Stolarski, R. & Shugar, D. (1983). Hydroxylamine mutagenesis: observation of inverted Watson-Crick base-pairing between N⁴-methoxycytosine and adenine with the aid of natural-abundance high-resolution ¹⁵N NMR spectroscopy. *Eur. J. Biochem.* **130**, 559-564.
- Patel, D. J., Kozlowski, S. A., Ikuta, S. & Itakura, K. (1984). Deoxyadenosine-deoxycytidine pairing in the d(C-G-C-G-A-A-T-T-C-A-C-G) duplex: conformation and dynamics at and adjacent to the dA·dC mismatch site. *Biochemistry*, **23**, 3218-3226.
- Quignard, E., Fazakerley, G. V., van der Marel, G., van Boom, J. H. & Guschlbauer, W. (1987). Comparison of the conformation of an oligonucleotide containing a central G-T base-pair with the non-mismatch sequence by proton NMR. *Nucl. Acids Res.* **15**, 3397-3409.
- Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H. & Kaptein, R. (1984). Sequential resonance assignments in ¹H NMR spectra of oligonucleotides by two-dimensional NMR spectroscopy. *Biochemistry*, **23**, 1371-1376.
- Sowers, L. C., Fazakerley, G. V., Kim, H., Dalton, L. & Goodman, M. F. (1986). Variation of nonexchangeable proton resonance chemical shifts as a probe of aberrant base-pair formation in DNA. *Biochemistry*, **25**, 3983-3988.
- Sowers, L. C., Shaw, B. R., Veigl, M. & Sedwick, D. W. (1987). DNA base modification: ionized base-pairs and mutagenesis. *Mutat. Res.* **177**, 201-218.
- Sowers, L. C., Eritja, R., Kaplan, B., Goodman, M. F. & Fazakerley, G. V. (1988). Equilibrium between a wobble and ionized base pair formed between fluorouracil and guanine in DNA as studied by proton and fluorine NMR. *J. Biol. Chem.* **263**, 14794-14801.
- Sowers, L. C., Goodman, M. F., Eritja, R., Kaplan, B. E. & Fazakerley, G. V. (1989). Ionized and wobble base-pairing for bromouracil-guanine in equilibrium under physiological conditions. *J. Mol. Biol.* **205**, 437-447.
- Spengler, S. & Singer, B. (1981). Effect of tautomeric shift on mutation: N⁴-methoxycytidine forms hydrogen bonds with adenosine in polymers. *Biochemistry*, **20**, 7290-7294.
- Stone, M. J., Nedderman, A. N. R., Williams, D. H., Lin, P. K. T. & Brown, D. (1991). Molecular basis for methoxyamine initiated mutagenesis. ¹H Nuclear magnetic resonance studies of base-modified oligonucleotides. *J. Mol. Biol.* **222**, 711-723.
- Topal, M. D. & Fresco, J. R. (1976). Complementary base pairing and the origin of substitution mutations. *Nature (London)*, **263**, 285-285.
- van Meervelt, L., Moore, M. H., Lin, P. K. T., Brown, D. M. & Kennard, O. (1990). Molecular and crystal structure of d(CGCGm⁴CG): N⁴-methoxycytosine-guanine base-pairs in Z-DNA. *J. Mol. Biol.* **216**, 773-781.
- Wang, C., Gao, H., Gaffney, B. L. & Jones, R. A. (1991). Nitrogen-15-labeled oligonucleotides. 3. Protonation of the adenine N1 in the A·C and A·G mispairs of the duplexes {d[CG(¹⁵N¹)AGAATTCCG]}₂ and {d[CGGAATTC(¹⁵N¹)ACG]}₂. *J. Amer. Chem. Soc.* **113**, 5486-5488.
- Watson, J. D. & Crick, F. H. C. (1953). Genetical implications of the structure of deoxyribonucleic acid. *Nature (London)*, **171**, 964-967.

II.

.II

Methoxyamine-induced mutagenesis of nucleic acids

proton NMR study of oligonucleotides containing *N*⁴-methoxycytosine paired with adenine or guanine

Elia GDANIEC¹, Buu BAN², Lawrence C. SOWERS³ and G. Victor FAZAKERLEY⁴

¹Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego, Poland
²Commissariat à l'Énergie Atomique, Service de Chimie Moléculaire, URA CNRS 331, CE Saclay, Gif-sur-Yvette, France

³Division of Pediatrics, City of Hope National Medical Center, Duarte, California, USA

⁴Commissariat à l'Énergie Atomique, Service de Biochimie et de Génétique Moléculaire, Département de Biologie Cellulaire et Moléculaire, CE Saclay, Gif-sur-Yvette Cedex, France

Received 27 June/9 September 1996) – EJB 96 0951/2

We report the solution structure of two heptanucleotides each containing a central *N*⁴-methoxycytosine, in one case paired with adenine on the opposite strand and the other with guanine. For the *N*⁴-methoxycytosine · adenine pair, only the imino form of the *N*⁴-methoxycytosine residue is observed and base pairing is in Watson-Crick geometry. However, rotation of the methoxy group about the N-OCH₃ bond is not constrained to a particular orientation although it must be *anti* to the N3 of *N*⁴-methoxycytosine. The slow exchange on a proton NMR time scale between the single strand and double strand forms is attributed to the strong preference of the *cis* conformation of the OCH₃ group in the single strand, which inhibits base pair formation. For the *N*⁴-methoxycytosine that is base paired with guanine, we observe an amino form in Watson-Crick geometry in slow exchange with a base paired imino form in wobble geometry. The amino form is predominant at low temperature whereas the imino form predominates above 313 K. We have measured the exchange rate between the two forms at 303 K and observed a value of approximately 1 s⁻¹. The relative ratio of amino and imino forms of *N*⁴-methoxycytosine is influenced by both the base that is in front and the temperature. Our results explain the preferential replacement of dTTP by *N*⁴-methoxycytosine in primer elongation.

Keywords: DNA; mutagenesis; tautomerism; NMR.

The reaction of methoxyamine with nucleic acids results in conversion of cytosine bases to *N*⁴-methoxycytosine (mo⁴C). Substitution of one hydrogen atom of the cytosine amino group by an electron-withdrawing methoxy group can induce formation of the imino tautomer form with tautomeric equilibrium constant orders of magnitude lower than those for the standard DNA bases.

Studies of the stability of oligonucleotide duplexes containing mo⁴C [1, 2] revealed that mo⁴C, which is ambivalent in its hydrogen bonding potential, forms stable base pairs with both adenine and guanine. It was predicted that mo⁴C would base pair with G in the amino form and with A in the imino form. An NMR study [3] proposed that mo⁴C when paired with A exists in a rapid equilibrium between Watson-Crick and wobble conformations.

An X-ray crystal structure [4] of an oligonucleotide containing a mo⁴C · G base pair, found that the base pair adopts a wobble conformation with the mo⁴C base in the imino form. The oligonucleotide was observed to be a Z DNA form. Recent NMR studies [5, 6] proposed wobble and Watson-Crick structures in equilibrium. In one of these studies [6], a further equilibrium between *syn* and *anti* conformations of the methoxy group

of mo⁴C · G in the imino form was invoked to explain the complexity of the NMR spectra.

We recently reported [5] preliminary NMR studies on two heptamer duplexes with mo⁴C incorporated into the central position with G or A on the opposite strand. On the basis of NMR spectra of exchangeable protons, we have suggested that when paired with A, mo⁴C is in the imino configuration and in Watson-Crick geometry. However, for the mo⁴C · G duplex two structurally distinct configurations were observed in equilibrium with each another. The analysis of exchangeable proton spectra showed that both imino and amino configurations of mo⁴C base pair with G. When mo⁴C is in the amino form, it pairs with G in Watson-Crick geometry. In the second species, mo⁴C was found in the imino configuration paired with G in wobble geometry.

In this paper, we present further NMR data that confirm and extend our previous structural assignments on mo⁴C base pairs in DNA. Based upon interactions between nonexchangeable protons, we have determined the influence of mo⁴C substitution on local and global helix structure. Further, examination of proton exchange kinetics allows us to determine the relative rate of exchange of the mo⁴C between different base-pairing environments.

Correspondence to G. V. Fazakerley, Commissariat à l'Énergie Atomique, Service de Biochimie et de Génétique Moléculaire, DBCM, CE Saclay, F-91191 Gif-sur-Yvette Cedex, France

Fax: +33 1 69 08 47 12.

Abbreviations. ROESY, rotating-frame Overhauser spectroscopy; 1D, one dimensional; 2D, two dimensional; 2QF, double quantum filtered correlated spectroscopy; mo⁴C, *N*⁴-methoxycytosine.

MATERIALS AND METHODS

*N*⁴-Methoxycytosine was incorporated into the oligonucleotides as previously described [5].

Duplex annealing was monitored by one dimensional (1D) NMR experiments. The appropriate pairs of oligonucleotides were heated to 353 K followed by slow cooling to form the following duplexes: 5'-d(C1-A2-G3-mo⁺C-G5-G6-C7) and 3'-d(G14-T13-C12-X11-C10-C9-G8), where (1) X11 = A11 or (2) X11 = G11. The strand concentrations of the duplexes were 4 mM dissolved in 10 mM sodium phosphate, 150 mM NaCl, and 0.2 mM EDTA.

NMR spectra were recorded on either AMX500 or AMX600 Bruker spectrometers in either 99.99% D₂O or 90% H₂O/10% D₂O.

NOESY spectra were recorded in the phase-sensitive mode [7] with mixing times of 50 ms and 400 ms for mo⁺C · A and 30 ms and 400 ms for mo⁺C · G. The NOESY spectra in H₂O were recorded with 150 ms or 200 ms mixing times. The time domain data sets consisted of 1024 points in the *t*₂ dimension and 256 or 512 increments in the *t*₁ dimension. After zero filling, the data were multiplied by a slightly shifted sine bell function in both dimensions except for the short mixing time experiments. For these, the data were multiplied by a $\pi/2$ shifted sine bell before Fourier transformation.

For spectra recorded in H₂O, the observation pulse was replaced by a jump and return sequence [8] and the pulse maximum was placed at 15 ppm. TOCSY spectra [9] were recorded in the phase-sensitive mode with 25 ms and 70 ms mixing times.

Pure absorption 2QF-COSY (double quantum filtered correlated spectroscopy) spectra [10–12] were obtained with a time-proportional phase incrementation scheme. 4 K data points in the *t*₂ dimension and 256 free induction decays were collected.

For the ROESY (rotating-frame Overhauser enhancement spectroscopy) spectra [13–14], a 200-ms spin lock was applied during the mixing period.

The chemical exchange study between the forms mo⁺Ca · G (mo⁺C base in the amino form) and mo⁺Ci · G (mo⁺C base in the imino form) was carried out at 303 K using a DANTE-Z sequence for selective resonance inversion [15].

The pulse sequence was,

$$(\theta_x - \tau_1 - \theta_{x-} - \tau_1)_n - \tau_2 - \pi/2 - \text{Acq} \pm,$$

where $\theta = 2.79$ ms, $\tau_1 = 100$ μ s, $\tau_2 =$ variable delay, $n = 300$, and with a 5-s delay between each pulse train to ensure return to equilibrium. Each experiment involved 22 values of the variable delay ranging from 30 ms to 2.2 s.

We use the letters A and B to describe the forms mo⁺Ca and mo⁺Ci, respectively. The magnetizations, M_x^A , M_x^B , M_z^+ and M_z^- when the site A is selectively inverted as a function of τ_2 are given by [16]:

$$M_x^A - M_x^A(\tau) = \frac{M_x^A}{\Delta} \{ (-k'_A + k'_B + \Delta)e^{\lambda_+ \tau} + (k'_A - k'_B + \Delta)e^{\lambda_- \tau} \} \quad (1)$$

and

$$M_x^B - M_x^B(\tau) = 2 \frac{k_B M_x^B}{\Delta} (e^{\lambda_+ \tau} - e^{\lambda_- \tau}) = 2 \frac{k_A M_x^A}{\Delta} (e^{\lambda_+ \tau} - e^{\lambda_- \tau}) \quad (2)$$

in which,

$$k_A M_x^A = k_B M_x^B$$

$$\lambda_{\pm} = -\frac{1}{2} (k'_A + k'_B) \pm \frac{1}{2} \Delta$$

and

$$k'_A = k_A + R_{1A}$$

k_A is the rate of exchange of A towards B, and R_{1A} is the spin lattice relaxation rate of A in the absence of exchange.

Further information was obtained from analysing the magnetization curves [Eqn (3) and Eqn (4)].

$$M_z^+(\tau) = M_z^+(\tau) + M_z^B(\tau) \quad (3)$$

$$M_z^-(\tau) = M_z^+(\tau) - M_z^B(\tau) \quad (4)$$

From Eqns (1–4), the initial slopes at the origin for the magnetization of site A, site B, their sum and their difference can be deduced [Eqn (5)]:

$$p^A(0) = -2 k'_A M_x^A$$

$$p^B(0) = 2 k_B M_x^B = 2 k_A M_x^A$$

$$p^+(0) = 2 R_{1A} M_x^A$$

$$p^-(0) = -2 (k'_A + k_A) M_x^A \quad (5)$$

RESULTS

Duplex mo⁺C · A. Assignment of the exchangeable proton resonances. The assignment of the exchangeable proton resonances was obtained from the analysis of a NOESY spectrum recorded with a mixing time of 150 ms at 274 K and pH 5.9. The spectrum was recorded at lower pH than those in D₂O (see below) to slow down the exchange with the bulk solvent. We monitored the non-exchangeable proton shifts as a function of pH, and observed no changes in the pH range 5–8 from which we can conclude that there is no conformational change in this pH range.

Three regions are shown in Fig. 1. The lower part shows imino-imino interactions. Starting from the only A · T base pair imino proton, we can follow the interbase imino connectivity through to the G5 imino proton. Although the resonances of the terminal G residues are strongly attenuated by exchange with solvent, the cross-peak corresponding to the interaction between the T13 and G14 imino protons is still visible. The remaining non-attenuated imino proton is assigned to G6, for which a weak cross-peak is observed with the C10 hydrogen bonded amino proton in the middle part of Fig. 1. This region corresponds to imino/amino/H2/H5 interactions and confirms the chain of connectivities observed for imino/imino interactions. Pairs of C and A amino cross-peaks are connected by solid lines.

The resonance at 12.01 ppm corresponds to an exchangeable proton of the mo⁺C · A base pair. This proton shows strong cross-peaks with the amino protons of A11. Observation of separate resonances for the amino protons of A11 strongly indicates that this group is involved in hydrogen bonding. When non-hydrogen bonded, an adenosine amino group rotates rapidly and gives rise to a single resonance. For A · T base pairs, the resonance of the hydrogen bonded proton is typically found in the range 7.7–8.1 ppm. For the mo⁺C base pair it is found at 8.49 ppm. This implies that hydrogen bonding is with a nitrogen acceptor. If the acceptor was an oxygen, as in Fig. 2a, we would expect a significant upfield shift relative to that observed for A · T base pairs. The resonance at 12.01 ppm also shows a very strong cross-peak with a non-exchangeable proton at 7.84 ppm, which must be the A11 H2. The intensity is typical for that of imino-H2 intrabase pair interactions of A · T pairs. This excludes the possibility of an amino form of mo⁺C in base pairing, Fig. 2b, for which the proton would be too far from A11 H2 to give such a cross-peak. Together with the observations for the A11 amino resonances, only one model, Fig. 2c, in which the mo⁺C · A base pair adopts Watson-Crick geometry with the mo⁺C base in the imino form, fits all the data. This indicates that the N4 methoxy group has to be *anti* relative to the N3 nitrogen because only the *anti* isomer of mo⁺C is able to partici-

These authors have shown that the
rate of reaction is first order in
the concentration of the reactant
and is independent of the concentration
of the other reactant. The rate
constant is $k = 0.1 \text{ min}^{-1}$. The
half-life of the reactant is 7
minutes.

For the reaction $A + B \rightarrow C$,
the rate of reaction is first order
in the concentration of A and
second order in the concentration
of B. The rate constant is $k = 0.01$
 min^{-1} . The half-life of A is 10
minutes.

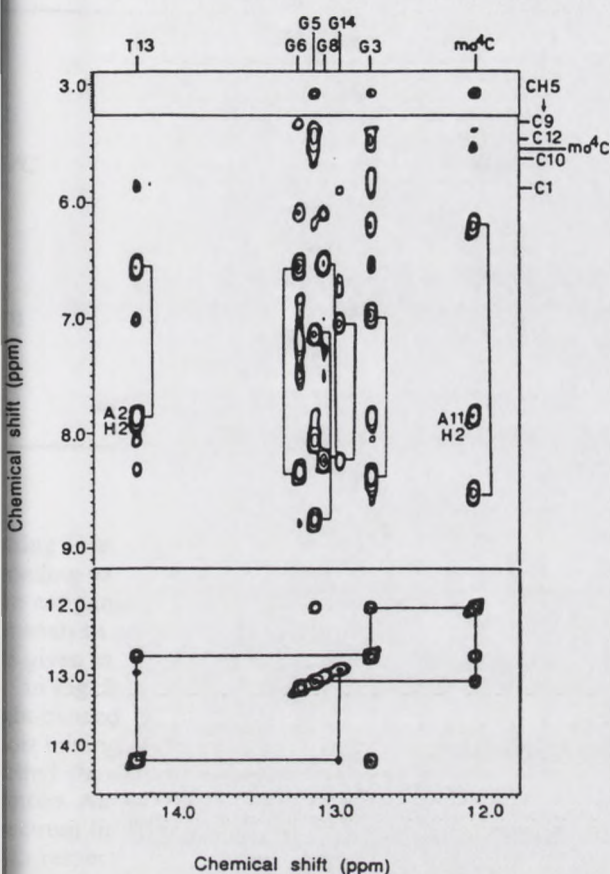
The overall order of the reaction
is 3. The rate of reaction is
proportional to the concentration
of A raised to the first power
and to the concentration of B
raised to the second power.

The rate of reaction is first order
in the concentration of A and
second order in the concentration
of B. The rate constant is $k = 0.01$
 min^{-1} . The half-life of A is 10
minutes.

The rate of reaction is first order
in the concentration of A and
second order in the concentration
of B. The rate constant is $k = 0.01$
 min^{-1} . The half-life of A is 10
minutes.

The rate of reaction is first order
in the concentration of A and
second order in the concentration
of B. The rate constant is $k = 0.01$
 min^{-1} . The half-life of A is 10
minutes.

The rate of reaction is first order
in the concentration of A and
second order in the concentration
of B. The rate constant is $k = 0.01$
 min^{-1} . The half-life of A is 10
minutes.



1. Three regions of the NOESY spectrum of the $mo^4C \cdot A$ duplex in 90% $H_2O/10\%$ D_2O recorded at pH 5.9 and 274 K with mixing time 150 ms. The assignment of the imino protons is shown on the horizontal axis. Lower part shows imino/amino interactions. The middle part shows interactions between imino and amino, CH5, AH2 protons. Pairs of C and A amino proton resonances are connected by solid lines for the intra-base interactions. The upper part shows interactions with the mo^4C methyl resonance.

interactions in hydrogen bonding with A by two hydrogen bonds as indicated from the data.

The upper part of Fig. 1 shows interactions with the mo^4C methyl group (assignment, see below). The methyl group shows NOEs (nuclear Overhauser effects) to the imino protons of the adjacent base pairs and to its own imino N3 proton resonance. Examination of the cross-section through the $mo^4C(OCH_3)$ resonance shows also strong NOEs to the amino protons of A11 and amino protons of both C10 and C12. The above results suggest that the methoxy group does not occupy a fixed position and may be orientated in either the 5' or the 3' direction.

Assignment of the non-exchangeable proton resonances. Assignment of the non-exchangeable proton resonances was obtained from analysis of the NOESY spectra recorded with 10 ms and 50 ms mixing times and the TOCSY spectra in D_2O . The region corresponding to interactions between the base H8/H2 protons and the H1'/H5 protons of the 400-ms NOESY spectrum recorded at 298 K is shown in Fig. 3. Six strong cross-peaks (marked X) correspond to interactions between H5-H6 protons of cytidine residues. Starting from the 5'-terminal cytidine at 7.65 ppm, the sequential connectivities can be followed without ambiguity up to C7. Similarly on the other strand, the connectivities can be followed from G8 to G14.

Analysis of the region of interactions between the base H8/H2 protons and H2'/H2''/CH₃ protons (data not shown) confirms assignment of the base protons shown in Fig. 3. The NOESY

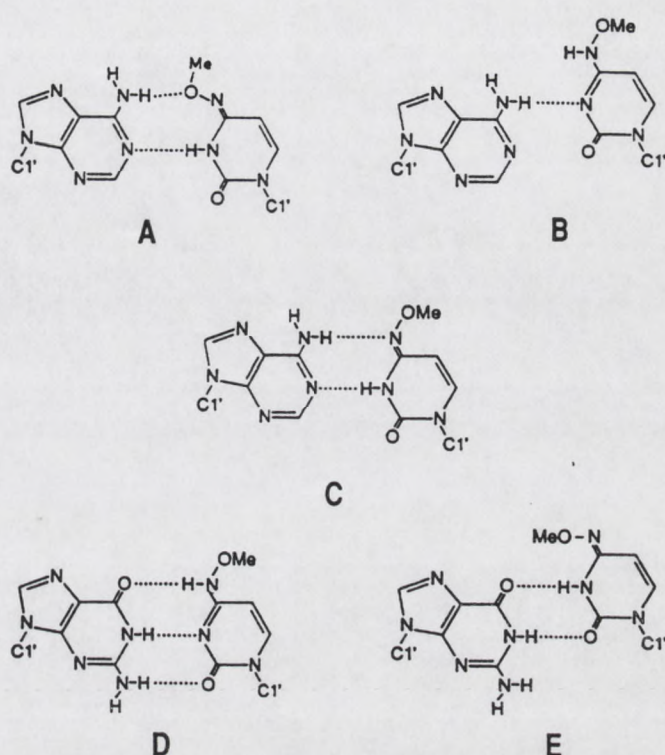


Fig. 2. Possible base pairing forms for $mo^4C \cdot A$ and $mo^4C \cdot G$.

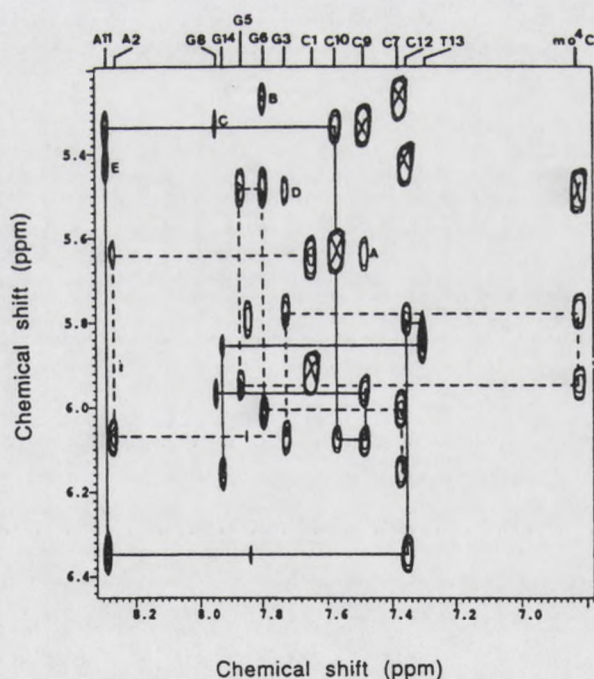


Fig. 3. Part of the 400-ms NOESY spectrum of the $mo^4C \cdot A$ duplex recorded at 298 K, pH 7. The cross-peaks marked with an X correspond to CH6-CH5 interactions. Interbase cross-peaks labelled A-E are described in the text.

experiment with a short mixing time allows an unambiguous discrimination between H2' and H2'' protons, as the H2''-H1' NOE is always larger than that for H2'-H1'.

The A2 H2 and A11 H2 resonances each exhibit a weak cross-peak with their own H1'. The NOEs of the H2 resonances with their adjacent G3 H1' and C12 H1' resonances overlap re-



The upper part of the graph shows a linear increase in Y from 0.0 to 0.5 as X increases from 0.0 to 0.5. The lower part of the graph shows a linear decrease in Y from 0.4 to 0.0 as X increases from 0.6 to 1.0. The overall trend is a peak at X=0.5, Y=0.5.

Table 1. Chemical shifts of non-exchangeable protons at 298 K and of exchangeable protons at 274 K for the mo⁴C · A duplex. The chemical shift of the methoxy methyl group of mo⁴C is 3.18 ppm.

Residue	Chemical shift for								
	H5/H6	H5/H2 CH ₃	H1'	H2'	H2''	H3'	H4'	NH	NH ₂
	ppm								
1	7.65	5.92	5.64	1.88	2.35	4.68	4.02	—	8.28/7.04
2	8.26	7.84	6.06	2.78	2.93	5.03	4.39	—	8.08/6.58
3	7.72	—	5.77	2.51	2.51	4.95	4.39	12.71	—
mo ⁴ C	6.82	5.49	5.94	1.97	2.41	4.92	4.25	12.01	—
5	7.86	—	5.48	2.65	2.65	4.98	4.31	13.09	8.07/5.40
5	7.79	—	6.01	2.54	2.72	4.97	4.41	13.18	—
7	7.38	5.27	6.14	2.20	2.20	4.50	4.02	—	8.26/6.52
8	7.94	—	5.96	2.64	2.75	4.82	4.25	13.01	—
8	7.48	5.33	6.07	2.12	2.47	4.83	4.25	—	8.35/6.56
10	7.57	5.63	5.33	2.20	2.39	4.86	4.12	—	8.78/7.15
11	8.28	7.83	6.33	2.76	2.98	5.06	4.46	—	8.49/6.19
12	7.35	5.43	5.79	1.91	2.40	4.76	4.34	—	8.39/6.99
13	7.30	1.65	5.86	1.96	2.34	4.85	4.11	14.22	—
14	7.92	—	6.15	2.63	2.38	4.68	4.17	12.92	—

...ing in a strong cross-peak at 7.83 ppm. Cross-peaks corresponding to interbase interactions H8/H6-H5 are labelled A-E. The assignment of the H3' and H4' resonances was obtained by the analysis of the TOCSY spectra. The observed chemical shifts are given in Table 1.

In Fig. 3, all inter- and intrasidue interactions expected for right-handed B DNA are observed. The cross-section of the short mixing time NOESY spectrum taken through the methoxy methyl group shows interactions with G3 H8, mo⁴C H5, and C10 H5 protons. All of them are weak and of similar intensity. From the spectrum in H₂O, we know that the N4 methoxy group is *anti* with respect to N3 and that the methyl group can point either towards the 3'- or 5'-direction.

The presence of the methoxy group apparently introduces some changes in the local geometry of the helix. We have checked, by model building, that in a normal B-helix the distance between the methyl group and the C10 H5 proton would be too long to observe any NOE effect. The only explanation for this interaction is that the mo⁴C · A base pair is displaced towards the major groove and the helical twist increases for the step G3 · C12-mo⁴C · A11. This arrangement of the base pairs in result in the interactions observed in the NOESY spectra.

To determine if the above distortion influences the sugar conformations, we have examined the relative intensity of the intrasidue cross-peaks between the base H6/H8 protons and the H2' versus H3' protons. This ratio gives a good indication of the sugar pucker [17] and we have found that for all non-terminal residues, within experimental error, it corresponds to a predominantly C2'-*endo* conformation. Additionally, these results were confirmed by analysis of the 2QF-COSY spectrum (data not shown), for which the sum of the JH1'-H2' and JH1'-H2'' coupling constants is greater than 14.6 Hz for all non-terminal residues.

The only significant deviation observed in the 50-ms mixing time NOESY spectrum from that of a normal B DNA structure is the ratio of the cross-peak volumes corresponding to interresidue interactions between the G5 H8 and mo⁴C H2', H2'' protons. This ratio is typically about 10 with short mixing times, except for the above interaction where the ratio is observed to be 1.5. This can be accounted for by an unusually small helical twist between the base pairs G5 · C10 and mo⁴C · A11 of 15–20 degrees to accommodate the methoxy group better.

Duplex mo⁴C · G. On the basis of NOESY spectra in 90% H₂O/10% D₂O, we have previously shown [5] that mo⁴C when paired with G exists in an equilibrium between imino and amino configurations. The amino and imino tautomers of the mo⁴C base pair with G in Watson-Crick and wobble geometry are shown in Fig. 2d and e, respectively.

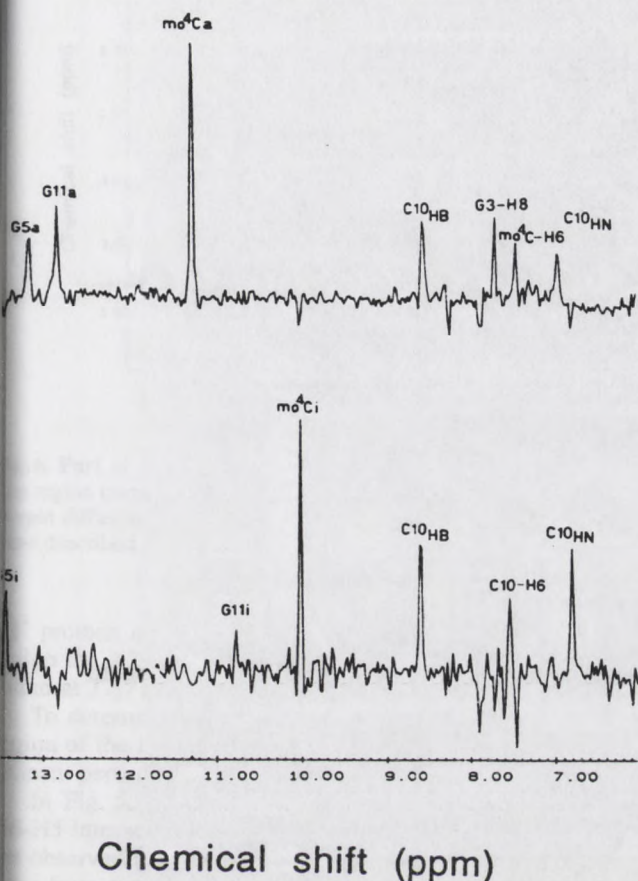
Complete assignment of the exchangeable protons based on their connectivities to other exchangeable and non-exchangeable protons has been carried out and the observed chemical shifts are given in Table 2.

N⁴-substituted cytidines generally have a preferred conformation for the substituent group *syn* relative to N3 [18–21]. In this conformation of the base, the amino form of mo⁴C could only form one hydrogen bond with guanine, between the G N1 imino proton and the mo⁴C C2 carbonyl. However, the imino form of the base can form two hydrogen bonds in which the bases are in wobble geometry, Fig. 2e, and this is what we observe. In this wobble geometry, no constraints are imposed on the methoxy group. We observe a strong NOE between the methyl group and the N3 imino proton of mo⁴C at 10.00 ppm confirming that the methoxy group remains *syn* (Fig. 4A). In the *anti* form, this NOE would not be expected. Additionally, we observe NOEs between the methyl group and the C10 amino protons. By spin diffusion, weak NOEs are also observed to the G11 and G5 imino protons.

For mo⁴C in the amino form, the methoxy group has to rotate *anti* to N3 for stable hydrogen bonding to occur. Three hydrogen bonds can be formed between the bases that are in Watson-Crick geometry (Fig. 2d). We observe NOEs from the mo⁴C methyl group to both amino protons of C10 and also to the G3 H8 proton (Fig. 4B), which suggests that there is rotational freedom about the N-OCH₃ bond and that the methyl group can be close to either the G3 · C12 or G5 · C10 base pairs.

Assignment of the non-exchangeable protons. NOESY spectra for the duplex mo⁴C · G were recorded at 283, 288, 293, and 303 K. Only for the spectra recorded at 283 K and 288 K could we follow the sequential H6/H8-H1' proton pathway without ambiguity. At these temperatures, all of the H1' and H6/H8 resonances could be identified.

Fig. 5 shows a region of the NOESY spectrum recorded at 283 K. For clarity, the sequential H8/H6-H1' assignments for the



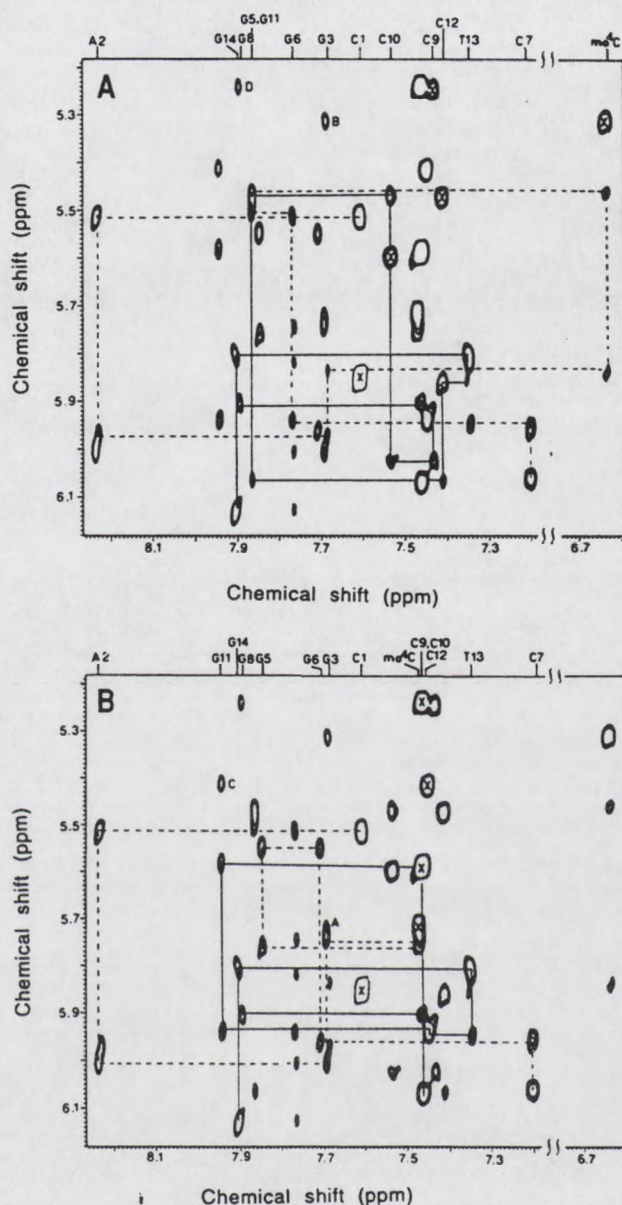
4. The cross-section of the 200-ms NOESY spectrum recorded in 10% $H_2O/10\%$ D_2O at pH 6.2 and 274 K. Rows are taken through methoxy methyl of the (A) imino and (B) amino forms of $mo^4C \cdot G$. Peaks HB and NH correspond to the hydrogen bonded and non-hydrogen bonded amino protons of C10, respectively.

ino form, $mo^4Ca \cdot G$ and the imino form, $mo^4Ci \cdot G$ are shown separately. On the basis of the TOCSY spectrum, we can assign the cross-peaks marked X to the intranucleotide CH5-CH6 interactions.

The most convenient starting point for the sequential assignment is from the resonance at 6.64 ppm, which can be assigned to mo^4Ci H6 on the basis of its chemical shift. Without any ambiguity we can follow the connectivities in the 5' direction to G11i (Fig. 5A). In the other direction, the connectivities are not unambiguous due to poor spectral resolution. This is resolved after an analysis of the region corresponding to the interactions $H8-H2'/H2''$ (data not shown).

At this temperature, the amino form is more populated than the imino form. This helps us to discriminate between these two forms as we expect the cross-peaks from the imino form to be of lower intensity. With this in mind, the connectivities on the other strand can be followed from G14i up to G11i. The ambiguity arising from the overlapping G11i H8 and G5i H8 resonances can be resolved referring again to the $H8/H6-H2'/H2''$ region and the sequential pathway can be traced out as is shown in Fig. 5A.

For the amino form, we start with the terminal 5'-C1a and follow the sequential connectivities up to G3a (Fig. 5B). From the other end of the strand, we can follow the chain up to mo^4Ca . The ambiguity in tracing the G3a- mo^4Ca -G5a chain arises from the overlapping of the interbase mo^4Ca H5-G3a H8 cross-peak with the G3a H8-G3a H1' cross-peak. Both the intra- and inter-nucleotide mo^4Ca H6-H1' NOEs also overlap. Fig. 6 shows the



5. Part of the 400-ms NOESY spectrum of $mo^4C \cdot G$ duplex recorded at 283 K. The region corresponds to interactions between base protons and $H1'/H5$ protons. Sequential assignment for the imino and amino form of $mo^4C \cdot G$ duplex are traced separately in A and B, respectively. Cross-peaks marked with an X correspond to CH6-CH5 interactions.

region corresponding to $H1'/H5-H2'/H2''$ interactions. This enables precise assignment of the chemical shifts of the $H1'$ resonances where overlap occurs. The resonances at 5.76, 5.74, and 5.71 ppm correspond to C4a $H1'$, G3a $H1'$, and C4a H5, respectively. After assignment of these protons, we can trace the pattern of sequential NOEs for this strand as shown in Fig. 5B.

For the other strand, the connectivities can be followed, with the aid of other regions of the spectrum, without interruption, although the intranucleotide C10 $H1'$ peak is coincident with the H5-H6 cross-peak and the H6 protons of C9 and C10 are coincident. We do not expect large chemical shift differences for the H5 protons of the C residues between the two forms. We can thus assign the CH5-CH6 cross-peak next to that of C9i to that of C9a and the other at 7.46 ppm to that of C10a.

Four cross-peaks are observed with a resonance at 7.77 ppm (Fig. 5), which correspond to interactions of A2a H2 with the



Chemical analysis

The analysis of the sample was performed using a standard method. The results are as follows:

The sample was found to contain the following components:

- Component A: 10.5%
- Component B: 25.3%
- Component C: 15.7%
- Component D: 8.2%
- Component E: 42.3%

The total percentage of the components is 100.0%.

The analysis was conducted on 10/25/2023.

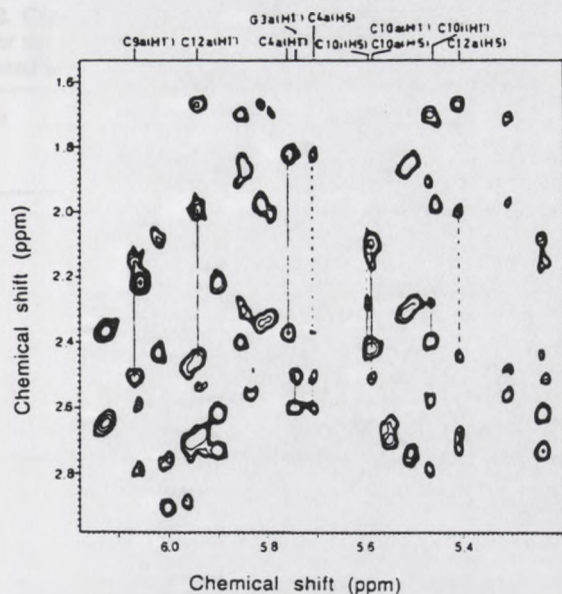


Fig. 6. Part of the 400-ms NOESY spectrum of the $mo^4C \cdot G$ duplex. The region corresponds to $H1'/H5-H2'/H2''$ interactions. Interactions due to spin diffusion are linked by broken lines. For clarity, only the interactions described in the text are indicated.

$H1'$ protons of G3a, G14a, T13a, and its own $H1'$. We cannot assign the A2i H2 on the basis of the NOESY spectrum. It was found at 7.87 ppm in the NOESY spectrum in H_2O .

To determine the stacking, we have examined the aromatic region of the NOESY spectrum and we have not found any significant perturbation in either duplex.

In Fig. 5, four interbase cross-peaks corresponding to H8/H6-H5 interactions, which are typical for right-handed B-DNA, are observed (peaks A–D). Three of these interbase cross-peaks arise from the central base pairs: C4a-G3a, C4i-G3i, and C12a-G11a, peaks A–C, respectively. The other corresponds to C9a,i-G8, peak D.

The $H2'$ and $H2''$ resonances were assigned from analysis of the NOESY spectrum recorded with a 30-ms mixing time. We were not able to assign all the $H4'$ resonances because of strong signal overlap.

Detailed analysis of 2QF-COSY spectra of the $mo^4C \cdot G$ duplex was impossible because of strong overlap of resonance. Some of signals from the imino form were also too low in intensity to give precise values of spin-spin coupling constants. However, we could measure the sum of $JH1'-H2'$ and $JH1'-H2''$ for the mo^4C residues in both forms. This was found to be greater than 14.6 Hz in both cases, which indicates a predominantly $C2'$ *endo* conformation. The ratio of the aromatic proton to $H2'/H3'$ NOE also gives a measure of the $C2'$ to $C3'$ *endo* conformations. With the exception of the C7 residue, the NOE with the $H2'$ proton was found to be very much larger than with the $H3'$ proton, which confirms the presence of predominantly $C2'$ *endo* sugar conformations.

We have measured the internucleotide NOEs in the 30-ms NOESY spectrum and we observe that these are all in agreement with a globally B form DNA in which the $mo^4C \cdot G$ bases are intrahelical and base paired.

Table 2 summarizes the observed chemical shifts for $mo^4Ci \cdot G$ and $mo^4Ca \cdot G$ duplexes.

The cross-section of the 30-ms mixing time NOESY spectrum taken through the methyl group of $mo^4Ci \cdot G$ shows only an interaction with the H5 proton of C10i. This interaction is possible only when the methoxy group adopts the *syn* confor-

mation relative to N3. The cross-section through the methyl group of mo^4Ca reveals NOEs to the C4a H5 and the G3a H8 protons. This is in agreement with the previous data, in which it was found that when mo^4C pairs with G in the amino form, the methoxy group is *anti*.

Additional cross-peaks were observed in the aromatic-aromatic and $H1'/H5-H1'/H5$ regions of NOESY spectra recorded above 293 K. We have recorded ROESY spectra to determine whether they arise from cross-relaxation or from chemical exchange. All these new cross-peaks are in phase with the diagonal and must therefore arise from chemical exchange between the two forms. Below 293 K, we do not observe cross-peaks from exchange between the $mo^4Ci \cdot G$ and $mo^4Ca \cdot G$ duplexes as the exchange is too slow.

The 1H -NMR spectra of the $mo^4C \cdot A$ and $mo^4C \cdot G$ duplexes as a function of temperature.

To determine the melting temperature of the $mo^4C \cdot A$ and $mo^4C \cdot G$ duplexes, we have recorded 1H -NMR spectra in D_2O as a function of temperature. For the $mo^4C \cdot A$ duplex, we were not able to follow the chemical shift changes of many of the proton resonances between 308 K and 328 K due to excessive line broadening. At temperatures above 308 K, many new resonances appeared, which indicates slow exchange between the helix and coil forms.

The mo^4C H6 proton resonance is well separated from the other base protons resonances. At 310 K, a new signal at 6.68 ppm appears (Fig. 7a), and was attributed to the mo^4C H6 coil resonance. To characterize the helix-coil transition, we have measured the ratio of integrals between helix and coil resonances for $mo^4C(H6)$ proton. We find that at 320 K these two species are equally populated. The profile for the methoxy protons is shown in Fig. 7b.

Two representative profiles of the chemical shift changes versus temperature for the $mo^4C \cdot G$ duplex are shown in Fig. 7c, d. All resonances of the imino form of $mo^4C \cdot G$ show typical sigmoidal profiles, which correspond to the double strand-single strand transition. The apparent T_m was defined as the mid-point of the sigmoidal melting curves and was 312 K. We can observe resonances of the amino form in Fig. 7c, d up to about 313 K, where they are of low intensity and also broad. At higher temperatures, the imino, amino, and coil forms are in rapid exchange.

We tried to measure the relative populations of the imino and amino forms by integrating the mo^4C amino and mo^4C imino protons. Above 308 K, both resonances were too broad to determine this ratio. At 318 K, we have observed only the signals from the imino form. Because of extensive overlapping of all of non-exchangeable protons, we could not determine this ratio in D_2O .

Kinetics. As the H6 resonance of the mo^4C base in the $mo^4Ci \cdot G$ duplex is well resolved, a totally selective inversion of the peak was possible. This experiment was performed at 303 K. Fig. 8 shows the change in intensity of the resonance inverted and that of the same proton in the duplex $mo^4Ca \cdot G$ as a function of time. From measurement of the initial slopes in Fig. 8 and using Eqns (5), estimates of the following parameters were obtained:

$$k'_A = 3.3 \text{ s}^{-1}, k_A = 1.1 \text{ s}^{-1}, R_{1A} = 2.2 \text{ s}^{-1}, \text{ and } k'_A + k_A = 4.5 \text{ s}^{-1}.$$

From these data, we have fitted the experimental data in Fig. 8 which gives the following expressions for Eqns (6),

$$M_2^A(\tau) = 105 e^{-1.6\tau} + 85 e^{-5.8\tau}$$

$$M_2^B(\tau) = 50 (e^{-1.6\tau} + e^{-5.8\tau}). \quad (6)$$



Figure 1. ¹H NMR spectrum of the compound in CDCl₃. The spectrum shows characteristic peaks for the structure, including a multiplet at 7.2 ppm and a singlet at 1.2 ppm.

The ¹H NMR spectrum of the compound in CDCl₃ is shown in Figure 1. The spectrum shows characteristic peaks for the structure, including a multiplet at 7.2 ppm and a singlet at 1.2 ppm.

The ¹³C NMR spectrum of the compound in CDCl₃ is shown in Figure 2. The spectrum shows characteristic peaks for the structure, including a multiplet at 160 ppm and a singlet at 10 ppm.

The IR spectrum of the compound is shown in Figure 3. The spectrum shows characteristic absorption bands for the structure, including a strong band at 1700 cm⁻¹ and a band at 3000 cm⁻¹.

The mass spectrum of the compound is shown in Figure 4. The spectrum shows characteristic peaks for the structure, including a base peak at m/z 100 and a peak at m/z 200.

The elemental analysis of the compound is shown in Table 1. The calculated and found values for C, H, and N are listed in the table.

The synthesis of the compound is described in the experimental section. The reaction conditions and workup procedure are detailed below.

The compound was synthesized from the starting materials in a two-step process. The first step involved the reaction of the starting materials under reflux for 24 hours.

The second step involved the purification of the crude product by column chromatography using silica gel and a gradient of ethyl acetate in hexanes.

The purified compound was then dried under high vacuum and stored in a desiccator until use. The yield of the compound was 85%.

The compound was characterized by ¹H NMR, ¹³C NMR, IR, and mass spectrometry. The molecular weight of the compound was determined to be 200 g/mol.

The compound was found to be stable under ambient conditions and soluble in common organic solvents such as dichloromethane, chloroform, and ethyl acetate.

The compound was used in a series of biological assays to evaluate its activity. The results of these assays are discussed in the following section.

The compound showed significant activity in the assays performed. The IC₅₀ values for the compound were determined to be in the range of 100-200 nM.

The mechanism of action of the compound is currently unknown and is the subject of ongoing research. Further studies are required to elucidate the exact mechanism.

The authors thank the funding agencies for their support of this work. We also thank the technical staff for their assistance in the laboratory.

Table 2. Chemical shifts of non-exchangeable protons at 283 K and of exchangeable protons at 274 K for $m^5C \cdot G$ duplex. The chemical shifts for the methoxy methyl group of m^5C are 3.29 ppm and 3.47 ppm for the amino and imino form, respectively. The non-assigned resonances are marked*).

Residue	Chemical shift for								
	H8/H6	H5/H2 CH3	H1'	H2'	H2Y	H3'	H4'	NH	NH2
	ppm								
a	7.61	5.86	5.52	1.87	2.30	4.66	4.02	—	8.18/7.00
a	8.24	7.77	6.01	2.76	2.90	5.02	4.37	—	8.00/6.51
i	8.23	7.77	5.97	2.76	2.89	5.02	4.37	—	7.88/7.62
a	7.70	—	5.74	2.50	2.60	4.97	4.39	12.85	—
i	7.69	—	5.83	2.48	2.55	4.94	*)	12.62	—
4Ca	7.47	5.71	5.76	1.83	2.37	4.78	4.18	11.25	—
4Ci	6.64	5.31	5.46	1.72	1.98	4.77	3.99	10.00	—
a	7.85	—	5.55	2.70	2.70	4.95	4.40	13.11	—
i	7.86	—	5.51	2.74	2.74	4.97	*)	13.40	—
a	7.71	—	5.96	2.49	2.71	4.96	*)	13.11	—
i	7.78	—	5.94	2.53	2.69	4.99	4.40	13.11	—
a	7.20	4.95	6.06	2.22	2.22	4.48	3.96	—	8.18/6.42
a	7.90	—	5.91	2.62	2.73	4.81	4.28	12.95	—
a	7.46	5.24	6.07	2.15	2.51	4.86	4.24	—	8.29/6.40
i	7.74	5.25	6.03	2.09	2.43	4.83	*)	—	8.29/6.40
Oa	7.46	5.58	5.58	2.10	2.41	4.86	4.40	—	8.56/6.97
Oi	7.54	5.59	5.47	2.28	2.39	4.86	*)	—	8.58/6.82
1a	7.95	—	5.94	2.67	2.67	4.50	*)	12.80	—
1i	7.88	—	6.06	2.58	2.79	4.99	*)	10.76	—
2a	7.45	5.41	5.95	2.00	2.46	4.77	4.20	—	8.29/6.75
2i	7.42	5.48	5.87	1.90	2.40	4.77	*)	—	8.27/6.95
3a	7.35	1.66	5.81	1.98	2.34	4.85	4.13	14.19	—
3i	7.35	1.69	5.81	1.98	2.34	4.85	4.13	14.22	—
4	7.91	—	6.13	2.64	2.37	4.69	4.18	12.95	—

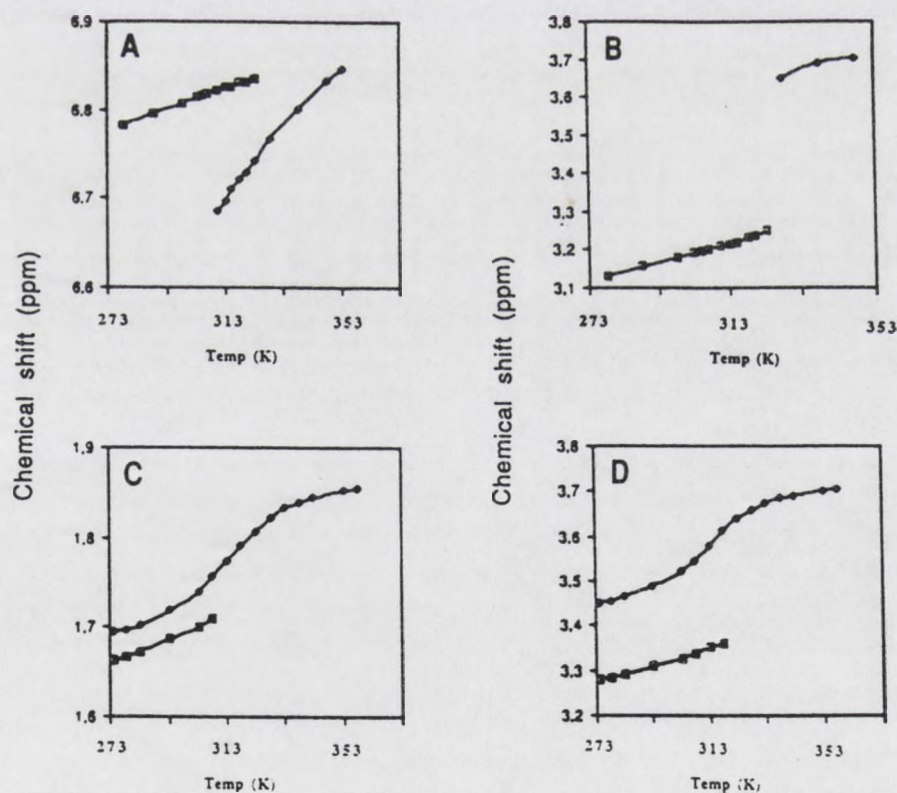


Fig. 7. Melting curves analysis. Melting curves for (a) the $m^5C(H6)$ proton and (b) the OCH_3 resonance of the $m^5C \cdot A$ duplex, (O) for the double strand form and (◆) for the single strand form. Similarly, (c) for the T13 (CH3) and (d) the OCH_3 resonance in the $m^5C \cdot G$ duplex where points marked (◆) correspond to the imino form which at temperatures above 313 K is in fast exchange with the single strand species and points marked (□), which correspond to the amino form.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----



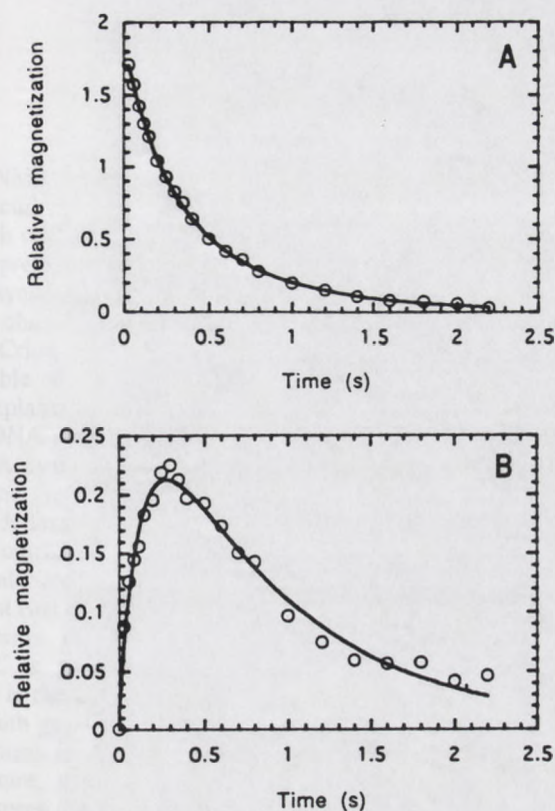


Fig. 8. Inversion and magnetization transfer for the $mo^4C(H6)$ proton. (A) After selective inversion of $mo^4C(H6) \cdot G$ as a function of time. (B) Resulting magnetization of the proton in the $mo^4Ca \cdot G$ duplex. The vertical scale is in relative units.

With the numerical values obtained from Eqns (6), we can calculate, using Eqns (1–3), the following kinetic parameters:

$$k'_A = 3.48 \text{ s}^{-1}, k'_B = 3.92 \text{ s}^{-1}, \text{ and } k_A = 1.11 \text{ s}^{-1}.$$

In the reverse experiment, selective inversion of the H6 resonance of $mo^4Ca \cdot G$, is not possible due to resonance overlap. However, we have observed that the two species have the same population at this temperature within $\pm 10\%$, which gives $k_A \approx k_B$. We were not able to study the exchange over a wide range of temperatures. At low temperature, the exchange is too slow and at higher temperature the system is too close to the T_m of the duplex.

DISCUSSION

Correct formation of base pairs during DNA replication is essential for maintaining the genetic code of all organisms and the occasional formation of base mismatches underlies the generation of both transition and transversion mutations. The ability of the normal bases to form unpreferred tautomeric configurations presents an opportunity to create incorrect base pairs during DNA replication, as originally proposed by Watson and Crick. Chemical modifications of DNA bases that increase the proportion of the incorrect tautomeric form would then be expected to increase the frequency of both base mismatch formation and genetic mutations.

Although the rare tautomer model for the generation of base substitution mutations has assumed enormous historical importance, supporting experimental evidence has been disturbingly absent. Within this context, base pairs containing cytosine residues modified with electron-withdrawing substituents, which in-

cludes methoxycytosine reported here, assume considerable significance. Methoxy- and hydroxylamine are known mutagens generating predominantly transition mutations *in vivo*. Reaction of these nucleophiles with DNA bases generates the corresponding cytosine derivatives. Of particular interest is that such substitutions are known to increase the proportion of the incorrect tautomeric form of the monomers in solution by several orders of magnitude. Because of the reduced energy for tautomerization in methoxycytosine, mismatches containing this modified base present an unusual opportunity to experimentally verify an essential, and as yet incomplete link, connecting unpreferred tautomer formation and mutagenesis. Preliminary data with this system from our laboratory and another laboratory support formation of base-pair-dependent tautomerization in DNA in solution. The present study confirms and extends previous data.

In contrast to cytosine, which prefers the amino tautomeric configuration by several orders of magnitude, methoxycytosine is found predominantly in the imino configuration in solution, although the energy difference between imino and amino forms is thought to be very low. When in the imino tautomeric form, mo^4C would be expected to pair with adenine. To date, both an X-ray crystallographic and two NMR studies have confirmed that mo^4C is indeed in a predominant imino configuration paired with adenine in Watson-Crick geometry in a duplex oligonucleotide. The current study confirms and extends these observations, and examination of the dynamic interchange of several possible structures provides new information on the phenomena of base mismatch formation and mutagenesis.

In addition to the requirement for the imino tautomeric form, formation of a base pair with adenine in Watson-Crick geometry also requires rotation of the methoxy group from a *syn* orientation relative to the N3 position, as observed for the monomer in solution, to an *anti* orientation. In this study, we have examined the $mo^4C \cdot$ adenine base pair in the central position of a seven-base-pair duplex. Our data indicate the formation of a Watson-Crick base pair in which the mo^4C is in the imino configuration and the methoxy group is *anti* relative to N3. All of the NOESY data show that the duplex adopts a regular B-form with only one interaction, that between the G5 H8 and the mo^4C H2'/H2'' protons being unusual. This may be due to a changed helical twist to accommodate the methoxy group better.

A previous report [3] suggested that this structure may be in rapid equilibrium with several other candidate non-Watson-Crick structures involving both *syn-anti* rotation of the methoxy group and/or imino/amino tautomerization. By contrast, our data strongly indicate that the Watson-Crick structure predominates and that such rapid equilibria for this base pair do not occur. We do not observe additional resonances nor line broadening that would indicate exchange between two or more conformations. All our data, in particular the magnitude of the NOEs recorded with short mixing times, are consistent with the presence of a unique Watson-Crick base pair conformation.

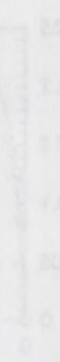
We do, however, observe strong NOEs from the methoxy group to protons on both adjacent base pairs. Although the *syn-anti* rotation of the methoxy group is constrained in this base pair, the C4-N-O-methyl torsion angle is not. As the methyl group could swing in either the 3' or the 5' direction, NOEs to both adjacent bases would be expected.

Previously, we have studied other systems in which the exocyclic substituents could interfere with formation of Watson-Crick base pairs such as N^4 -methylcytosine [21] and N^6 -methyladenine [22, 23]. As with mo^4C , the preferred position of the methyl substituent is *syn* to the hydrogen bonding face for the monomer in solution. To form Watson-Crick pairs, rotation must occur from the *syn* to *anti* orientation, and such rotation is apparently rate limiting. In such cases, one observes slow exchange.

relative humidity



relative humidity



A further increase in relative humidity is observed when the temperature is raised from 20°C to 30°C. This is due to the fact that the rate of evaporation is higher at the higher temperature.

In the present case, the relative humidity is 100% at 20°C and 100% at 30°C. This is due to the fact that the rate of evaporation is higher at the higher temperature.

The relative humidity is 100% at 20°C and 100% at 30°C. This is due to the fact that the rate of evaporation is higher at the higher temperature.

DISCUSSION

It is well known that the relative humidity of air is a function of temperature and pressure. In the present case, the relative humidity is 100% at 20°C and 100% at 30°C. This is due to the fact that the rate of evaporation is higher at the higher temperature.

Although the relative humidity is 100% at 20°C and 100% at 30°C, the rate of evaporation is higher at the higher temperature. This is due to the fact that the rate of evaporation is higher at the higher temperature.

the NMR time scale, between duplex and coil forms. The melting curves for the $\text{mo}^4\text{C} \cdot \text{adenine}$ duplex (Fig. 7), are in accord with such slow exchange and further increase the contrast-interpretation between our data and that published previously on the dynamics of the $\text{mo}^4\text{C} \cdot \text{adenine}$ base pair.

The observation of a rate-limiting step for formation of a Watson-Crick base pair between mo^4C and adenine, which is attributable to the *syn-anti* rotation of the *N*-methoxy group, explains why mo^4C deoxynucleotides substitute for dTTP during DNA synthesis but with considerably reduced efficiency. Although both mo^4C and T are predominantly in the imino tautomeric form, mo^4dCTP replaces dTTP with 2–3 orders of magnitude less efficiency.

In contrast to a single, preferred configuration for the $\text{C} \cdot \text{adenine}$ base pair, the $\text{mo}^4\text{C} \cdot \text{guanine}$ base pair is present in at least two configurations, which represent two distinct tautomeric forms. In one configuration, the amino tautomeric form of mo^4C is paired with guanine in Watson-Crick geometry whereas in the other configuration, the imino tautomeric form is paired with guanine in wobble geometry. The ratio of the amino/imino forms is observed to be temperature dependent. At low temperature, the amino form predominates whereas at higher temperatures the imino form predominates.

At intermediate temperatures, the two forms are in slow exchange with one another. Separate sets of resonances are observed for each configuration, including the H5, H6, and methyl exchangeable protons of mo^4C as well as the exchangeable amino/imino proton of mo^4C and the imino proton of the paired adenine residue. Distinct sets of NOEs are observed with the methyl protons and the adjacent bases when in either the wobble or *anti* (Watson-Crick) orientations. In a previous report [6], the methoxy protons were not assigned and therefore the position of the methoxy group could not be experimentally established. The *syn* orientation (Fig. 2e) reported here is the same as that determined previously by X-ray crystallography [4].

As expected for species that are in slow exchange with one another, the *syn* (wobble) and *anti* (Watson-Crick) forms display different melting curves as shown in Fig. 7. A typical sigmoidal curve is observed for the $\text{mo}^4\text{C} \cdot \text{G}$ duplex. In contrast to $\text{C} \cdot \text{A}$, the $\text{mo}^4\text{C} \cdot \text{G}$ pair is not in slow exchange with the *syn-anti* form as no *syn-anti* rotation is required for formation of the $\text{C} \cdot \text{G}$ wobble pair. Formation of the $\text{mo}^4\text{C} \cdot \text{G}$ pair does, however, require *syn-anti* rotation and thus the duplex and coil forms are observed in slow exchange. To form the $\text{mo}^4\text{C} \cdot \text{G}$ pair, two distinct conformational shifts must occur, specifically *syn-anti* rotation of the methoxy group and imino/amino tautomerization. It is as yet unknown if one step must occur before the other or which of the two steps is rate limiting.

We have measured the kinetics of exchange between the *syn-anti* forms at 303 K. Our data fit a two-state model well with $t_{1/2} \approx 1$ s. In contrast, a previous report [6] invoked the presence of at least three species in slow exchange to explain their data. We note that their system was comprised of an autocatalytic eight-residue duplex in which the duplex contains multiple mispairs. Our 2D (two dimensional) NMR experiments with two-residue duplexes containing a single, central mispair, were conducted at temperatures well below the T_m of our duplex and therefore the contribution from the single strand species is negligible. Fortunately, in our system for the $\text{mo}^4\text{C} \cdot \text{G}$ base pair, we observe two distinct species that are clearly defined and in slow exchange. Data reported previously were obtained at the T_m of the duplex.

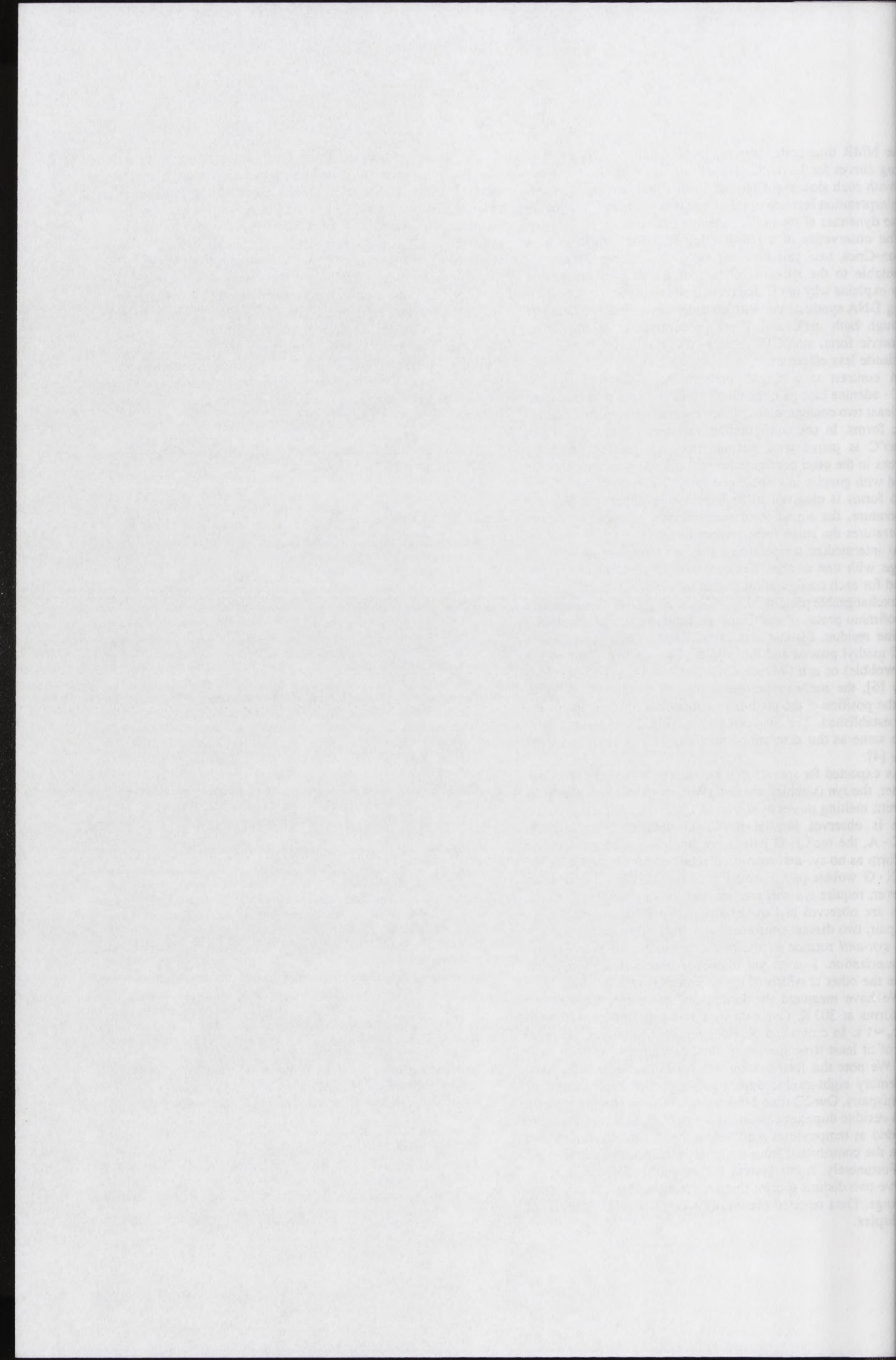
The data reported here strongly support previous suggestions that the ability of mo^4C residues to induce transition mutations results from its ability to form base pairs with both adenine and guanine. To our knowledge, mo^4C is the only case where two distinct tautomeric forms of a base (Fig. 2c and d) have been observed in DNA in Watson-Crick geometry in which the tautomeric form was determined by the opposing base in the helix.

Although preferentially in the imino tautomeric form, mo^4dCTP is only a poor replacement for dTTP, which can be attributed to the requirement for the methoxy group to rotate from the *syn* to *anti* conformation prior to formation of a Watson-Crick base pair. Substitution of mo^4dCTP for dCTP occurs with even lower efficiency [24]. In addition to rotation of the methoxy group from the *syn* to *anti* conformation, mo^4C must also undergo a second unpreferred equilibrium shift from the imino to the amino tautomeric form.

Support from the National Institutes of Health grants GM41336 and CA33572 is gratefully acknowledged.

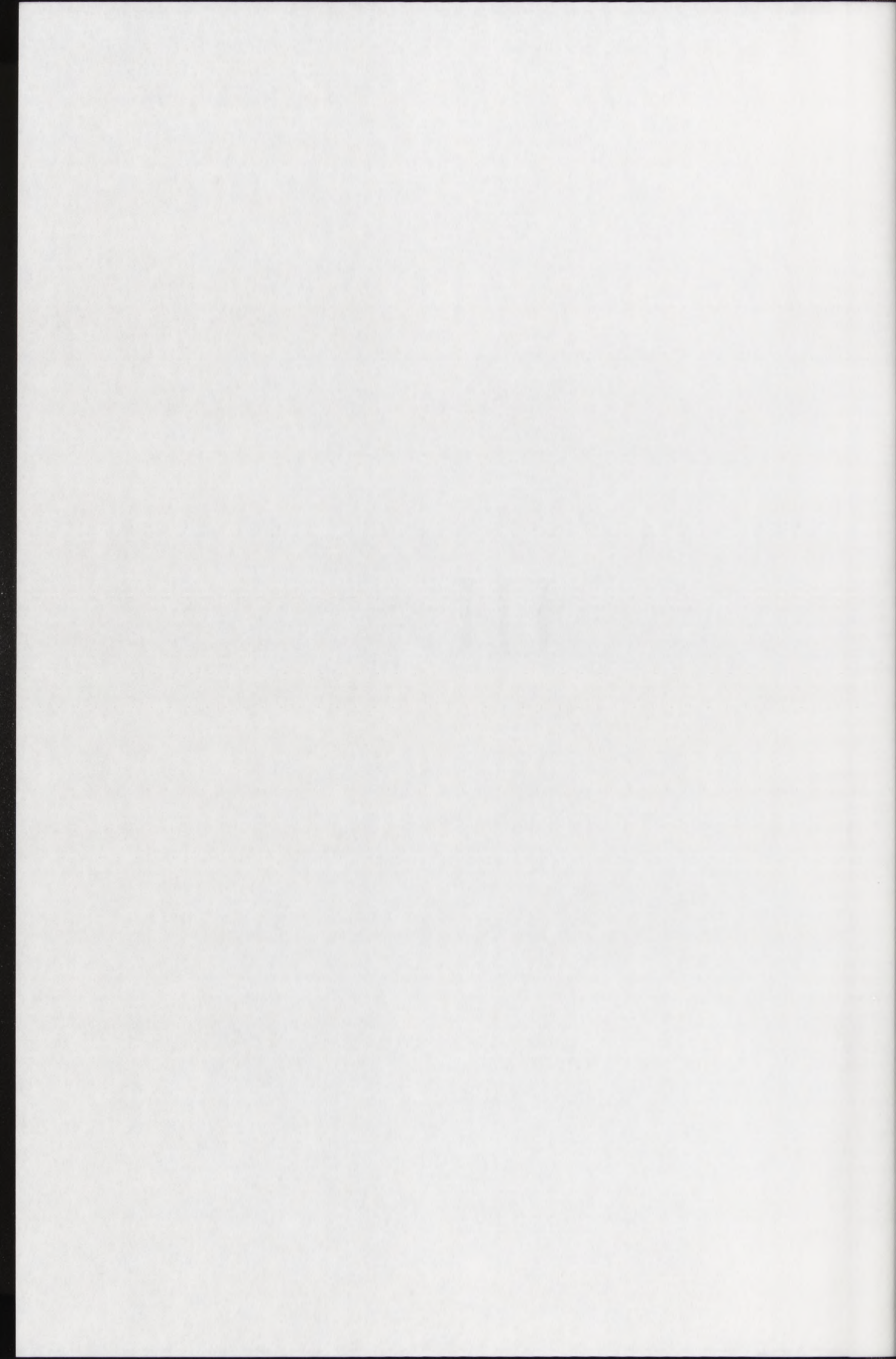
REFERENCES

- Anand, N. N., Brown, D. M. & Salisbury, S. A. (1987) *Nucleic Acids Res.* **15**, 8167–8176.
- Kong Thoo Lin, P. & Brown, D. M. (1989) *Nucleic Acids Res.* **17**, 10373–10383.
- Stone, M. J., Nedderman, A. N. R., Williams, D. H., Kong Thoo Lin, P. & Brown, D. M. (1991) *J. Mol. Biol.* **222**, 711–723.
- Van Meervelt, L., Moore, M. H., Kong Thoo Lin, P., Brown, D. M. & Kennard, O. (1990) *J. Mol. Biol.* **216**, 773–781.
- Fazakerley, G. V., Gdaniec, Z. & Sowers, L. C. (1993) *J. Mol. Biol.* **230**, 6–10.
- Nedderman, A. N. R., Stone, M. J., Williams, D. H., Kong Thoo Lin, P. & Brown, D. M. (1993) *J. Mol. Biol.* **230**, 1068–1076.
- Bodenhausen, G., Kogler, H. & Ernst, E. E. (1984) *J. Magn. Reson.* **58**, 370–338.
- Plateau, P. & Guéron, M. (1982) *J. Am. Chem. Soc.* **104**, 7310–7311.
- Davis, G. G. & Bax, A. (1985) *J. Am. Chem. Soc.* **107**, 2820–2821.
- Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wüthrich, K. (1984) *Biochem. Biophys. Res. Commun.* **117**, 479–485.
- Shaka, A. J. & Freeman, R. (1983) *J. Magn. Reson.* **51**, 169–173.
- Piantini, O. W., Sorensen, O. W. & Ernst, R. R. (1982) *J. Am. Chem. Soc.* **104**, 6800–6801.
- Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D. & Jeanolz, R. W. (1984) *J. Am. Chem. Soc.* **106**, 811–813.
- Bax, A. & Davis, D. G. J. (1985) *J. Magn. Reson.* **63**, 207–213.
- Boudot, D., Canet, D., Brondeau, J. & Boubel, J. C. (1989) *J. Magn. Reson.* **83**, 428–439.
- Hofman, R. A. & Forsén, S. (1966) *J. Chem. Phys.* **45**, 2049–2060.
- Cuniasse, P., Sowers, L. C., Eritja, A., Kaplan, B., Goodman, M. F., Cognet, J. A. H., Le Bret, M., Guschlbauer, W. & Fazakerley, G. V. (1989) *Biochemistry* **28**, 2018–2026.
- Brown, D. M., Hewlins, M. J. E. & Schell, P. (1968) *J. Chem. Soc. Sect. C Org. Chem.*, 1925–1929.
- Shugar, D., Huber, C. P. & Birnbaum, G. I. (1976) *Biochem. Biophys. Acta* **447**, 274–284.
- Engel, J. D. & von Hippel, P. H. (1974) *Biochemistry* **13**, 4143–4158.
- Fazakerley, G. V., Kraszewski, A., Téoule, R. & Guschlbauer, W. (1987) *Nucleic Acids Res.* **15**, 2191–2201.
- Fazakerley, G. V., Téoule, R., Guy, A., Fritzsche, H. & Guschlbauer, W. (1985) *Biochemistry* **24**, 4540–4548.
- Quignard, E., Fazakerley, G. V., Téoule, R., Guy, A. & Guschlbauer, W. (1985) *Eur. J. Biochem.* **152**, 99–105.
- Reeves, S. T. & Beattie, K. L. (1985) *Biochemistry* **24**, 2262–2268.



III.

RNA
regulation
is a field
Special
regula-
TAN
Budget
distur-
construc-
model
element
The
DNA pro-
animal
are ha-
loop to
a non-
refe- J
micro-
RNA
some I
factor
regions
regula-
rapid
The
Lan-
Center
MIB-8
The
1999
124



Iron Regulatory Element and Internal Loop/Bulge Structure for Ferritin mRNA Studied by Cobalt(III) Hexamine Binding, Molecular Modeling, and NMR Spectroscopy[†]

Zofia Gdaniec,^{‡,§} Hanna Sierzputowska-Gracz,^{§,||} and Elizabeth C. Theil^{*,§}

Departments of Biochemistry and Chemistry, North Carolina State University, Raleigh, North Carolina 27696-7622, and Institute for Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland

Received August 11, 1997

ABSTRACT: The ferritin IRE, a highly conserved (96–99% in vertebrates) mRNA translation regulatory element in animal mRNA, was studied by molecular modeling (using MC-SYM and DOCKING) and by NMR spectroscopy. Cobalt(III) hexamine was used to model hydrated Mg²⁺. IRE isoforms in other mRNAs regulate mRNA translation or stability; all IREs bind IRPs (iron regulatory proteins). A G·C base pair, conserved in ferritin IREs, spans an internal loop/bulge in the middle of an A-helix and, combined with a dynamic G·U base pair, formed a pocket suitable for Co(III) hexamine binding. On the basis of the effects of Co(III) hexamine on the ¹H NMR spectrum and results of automatic docking into the IRE model, the IRE bound Co(III) hexamine at the pocket in the major groove; Mg²⁺ may bind to the IRE at the same site on the basis of an analogy to Co(III) hexamine and on the Mg²⁺ inhibition of Cu(phen)₂ cleavage at the site. Distortion of the IRE helix by the internal loop/bulge near a conserved unpaired C required for IRP binding and adjacent to an IRP cross-linking site suggests a role for the pocket in ferritin IRE/IRP interactions.

RNA sequences in the noncoding region of mRNAs can regulate mRNA function. The predicted secondary structure is a hairpin distorted by a bulge, bulge loop, or internal loop. Specificity of the three-dimensional structure of RNA regulatory elements is recognized by proteins as in the Tat/TAR and Rev/RRE interactions of the HIV virus (1–3). Bulge loops and internal loops in RNA induce bends or distortions in helices, creating specific three-dimensional structures and, often, metal binding sites (4). Little is known about the three-dimensional structure of natural regulatory elements in eukaryotic mRNAs.

The IRE (iron responsive element) family of isoelements is a particularly well characterized control element in normal animal mRNAs encoding proteins of iron metabolism. IREs are hairpins of 9 or 10 base pairs, interrupted by a bulge loop of 1–4 nucleotides with a conserved C residue and with a terminal hexaloop, CAGUGX (reviewed most recently in refs 5–8). The metal complex Cu(phen)₂ binds at the internal bulge/loop (9, 10). All IREs recognize a family of RNA binding proteins, the IRPs (iron regulatory proteins); some IREs recognize other proteins as well, such as initiation factors (11, 12, 14). Single-copy IREs in the 5′-untranslated regions of mRNAs regulate ribosome binding, while pentuple-copy IREs in the 3′-untranslated regions are part of a rapid turnover element regulating mRNA stability; each type

of IRE is highly conserved (96–99%) which contrasts with the lower sequence conservation (35–45%) between translation and rapid turnover IREs (8).

The ferritin IRE is the best characterized IRE in terms of structure and function. Assurance of the biological relevance of IRE studies with synthetic RNA, used here and in other types of experiments, has been uniquely provided by earlier investigations using natural ferritin mRNA [poly(A⁺) RNA]¹ to study IRE structure, the IRP binding site and IRE function in regulating protein synthesis (9–11, 14–16); ferritin poly(A⁺) RNA showed function and/or chemical and enzymatic reactivity similar to those of the synthetic RNAs. The ferritin IRE is the most efficient of the translational regulatory IREs (13), possibly because of a conserved internal loop/bulge involving UGC/C rather than the bulge C of other IREs.

Previous NMR studies have focused on the role of the ferritin IRE terminal hexaloop (17, 18). In this study, a model of the complete IRE 30-mer is developed, assisted by NMR data from ¹⁵N- and ¹³C-labeled RNA and cobalt(III) hexamine/RNA complexes; the model is consistent with previous chemical and enzymatic studies. Co(III) hexamine significantly shifted proton NMR resonances of G7 and G27 in the internal loop/bulge region and docked in a pocket caused by distortion of the major groove in the middle of the IRE. The same region is also hypersensitive to cleavage by hydroxyl radical (16) and displays Mg-

[†] The work was supported in part by NIH Grant DK-20251.

* Corresponding author at Department of Biochemistry, North Carolina State University, Raleigh, NC 27695-7622. Phone: 919-515-5805. Fax: 919-515-5805. E-mail: Theil@bchserver.bch.ncsu.edu.

[‡] Polish Academy of Sciences.

[§] Department of Biochemistry, North Carolina State University.

^{||} Department of Chemistry, North Carolina State University.

¹ Poly(A⁺) RNA from a natural cell rich in ferritin mRNA [the embryonic red cell in which ~10% of the mRNA is ferritin mRNA (11)] was used with immunoprecipitation to examine control of ferritin synthesis (11, 14, 15) or with specific primers to examine IRE structure in the RNA after reaction with structure probes or IRP binding (9, 16).

Two Regions of the ...
Structure of the ...

The ...

The ...

The ...

The ...

The ...

The ...

The ...

The ...

The ...

The ...

The ...

The ...

The ...

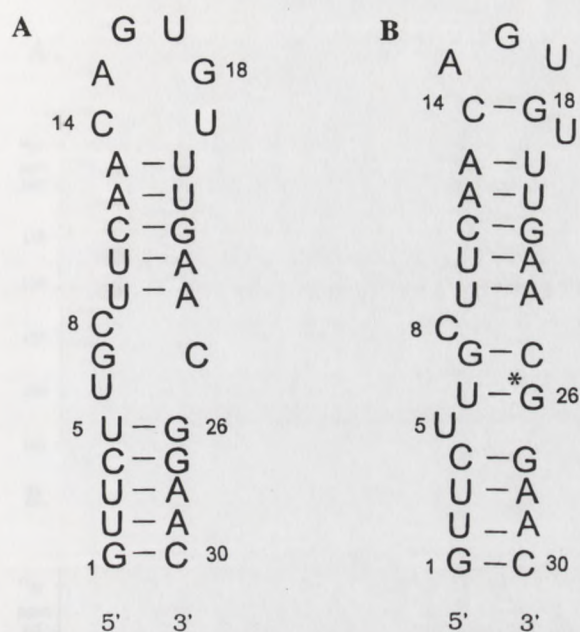


FIGURE 1: Secondary structure of the ferritin IRE. (A) The secondary structure predicted by m-fold (University of Wisconsin GCG-9.0); note the absence of the G-C base pairs across the hairpin loop and internal loop/bulge. (B) The secondary structure based on NMR spectroscopy (Figure 2). The asterisk indicates the dynamic, pH-sensitive G-U base pair (see Figure 2C).

sensitive changes in cleavage by $\text{Cu}(\text{phen})_2$ (9), indicating solvent accessibility and suggesting that hydrated Mg^{2+} binds at the site. In the IRE model, G7/U6 in the internal loop/bulge and G18/U19 which cross-link to the IRP (21) are 22 Å apart, in contrast to only 18 Å in an IRE model without the interhelical pocket, which may relate to correct positioning in the IRP binding site.

MATERIALS AND METHODS

RNA Synthesis. The 30-mer representing the frog ferritin IRE (Figure 1) was synthesized using the double-stranded T7 polymerase site and the complement of the 30-mer as a template, as previously described (17); vertebrate ferritin IREs are highly conserved (5–8) (96–99%), but the frog ferritin IRE is the only one which has been studied in natural [poly(A⁺)] mRNA as well as in synthetic mRNAs and RNA oligomers. The use of full-length double-stranded template increased the yield ~1.5–2-fold; reaction volumes were 24–90 mL. Cloned T7 polymerase was isolated as described by Studier et al. (22). RNA was purified by electrophoresis in urea/acrylamide gels as before (17), electroeluted, and concentrated by alcohol precipitation. To study the effect of pH on the detection of the G-U base pair (Figure 2C), commercially prepared (Cybersyn) RNA was used, but was purified by gel electrophoresis and dialyzed extensively against water before use.

RNA enriched in ^{13}C and ^{15}N was prepared using $^{13}\text{C}/^{15}\text{N}$ nucleotide triphosphates (NTPs) as described for the synthesis of RNA with natural abundance levels of the isotopes (19, 20). The ^{13}C - and ^{15}N -enriched NTPs were prepared using crude rRNA from *Methylophilus methylotrophus* provided by the NIH Research Resource for Heavy Atoms at Los Alamos National Laboratory. The crude rRNA was digested with

DNase, extracted with phenol and chloroform/isoamyl alcohol (24:1), and precipitated with alcohol followed by digestion to nucleotides with nuclease P1 (23) and conversion to NTPs using nucleoside monophosphate kinase, guanylate kinase, pyruvate kinase, myokinase, phosphoenolpyruvate, and ATP as described by Nikonowicz et al. (24). After concentration, lyophilization, and alcohol precipitation, the crude NTPs were dissolved in cold 1 M triethylamine/borate buffer (TEAB) at pH 9.5 and desalted on an Affigel 601 (Biorad) column equilibrated in 1 M TEAB buffer at 5 °C (25); the NTPs were eluted with cold distilled water acidified to pH 4–5 with CO_2 , lyophilized, dissolved in water, filtered through a washed Centricon 10 filter, and stored at pH 8.1 and –20 °C until use.

NMR Spectroscopy. RNA (0.5–1.0 mM in 10 mM sodium phosphate buffer and 0.1 mM EDTA at pH 6.8) was heated at 85 °C and slowly cooled in the NMR tube. Spectra were acquired on a Bruker DRX 500 MHz spectrometer. Spectra in H_2O were obtained either by the Watergate method (26) or by presaturation of the HDO signal for 2 s prior to applying an observation pulse or by using the jump–return water suppression and excitation maximum set to the imino resonances (27). Data for the two-dimensional (2D) NOESY experiment in 10% $\text{D}_2\text{O}/90\%$ H_2O were acquired at 12 °C using Watergate-water suppression [a 3–9–19 pulse sequence with the gradients for water suppression with excitation maximum set to the imino resonances (26)]. The spectrum was 2048 × 256 complex data points with a sweep width of 12 000 Hz, a mixing time of 250 ms, a recycle delay of 1.7 s, and 256 scans per slice. Spectra were processed with FELIX 95.0 software (Biosym/Molecular Simulations, Inc.) using an exponential weighting function or shifted sine-bell function to resolve overlapped imino protons.

NOESY, DQF-COSY, and TOCSY experiments were recorded in 99.996% D_2O on a 500 MHz GE Omega spectrometer or a Bruker 500 MHz spectrometer. Data sets with 2048 complex points in t_2 and 512 complex points in t_1 were acquired with 5000 Hz sweep widths in both dimensions and 128 scans per slice. NOESY spectra were acquired with mixing times of 120, 200, and 400 ms and a recycle delay of 2 s at 12 and 20 °C. The TOCSY spectrum was recorded with a 75 ms MLEV spin lock pulse and a recycle delay of 1.5 s. The DQF-COSY spectra were recorded with WALTZ decoupling of ^{31}P during acquisition and a recycle delay of 1.6 s. The diagonal and cross-peaks of DQF-COSY spectra were phased with antiphase absorption line shape in both directions. All spectra were processed with combinations of exponential and sine-skewed functions and zero-filled to 2K × 2K data points using XWINNMR Bruker or Felix 95.0 software.

Spectra with Co(III) hexammine and with various pHs were acquired on a Bruker DRX 500 MHz system. Imino proton spectra were obtained by the Watergate method (26). Typically, 2048 scans were collected. ^1H spectra in 10% $\text{D}_2\text{O}/90\%$ H_2O were collected at 12 °C in 16K point data sets consisting of 1024 scans each.

Spectra of double-labeled RNA were obtained on a Varian Unity Plus 600 MHz NMR spectrometer at the University of Chicago, Biological Sciences Division NMR Facility, used in consultation with Dr. Klaas Hallenga. The 2D (^1H – ^{15}N) HSQC experiments were carried out using gradients during the INEPT and reverse INEPT magnetization transfer delays

DMAP reacted with glycol and chloroformate. Glycol (2.0 g) and pyridine were added followed by sodium hydroxide with sodium PI (25) and then sodium hydroxide with sodium PI (25) and then sodium hydroxide with sodium PI (25) and then sodium hydroxide with sodium PI (25). After the reaction, the mixture was poured into water and extracted with diethyl ether. The ether extract was washed with water, dried over anhydrous calcium chloride, and then evaporated. The residue was purified by silica gel chromatography. The yield was 0.5 g (25%).

For synthesis of 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 100 mg of 1,2-dipalmitoyl-sn-glycero-3-phosphate (DPP) was dissolved in 10 mL of chloroform. To this was added 100 mg of 1,2-dipalmitoyl-sn-glycero-3-phosphate (DPP) and 100 mg of 1,2-dipalmitoyl-sn-glycero-3-phosphate (DPP). The mixture was stirred for 24 hours at room temperature. The mixture was then washed with water, dried over anhydrous calcium chloride, and then evaporated. The residue was purified by silica gel chromatography. The yield was 100 mg (100%).

For synthesis of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPPG), 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP) was dissolved in 10 mL of chloroform. To this was added 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP) and 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP). The mixture was stirred for 24 hours at room temperature. The mixture was then washed with water, dried over anhydrous calcium chloride, and then evaporated. The residue was purified by silica gel chromatography. The yield was 100 mg (100%).

For synthesis of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPPG), 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP) was dissolved in 10 mL of chloroform. To this was added 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP) and 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP). The mixture was stirred for 24 hours at room temperature. The mixture was then washed with water, dried over anhydrous calcium chloride, and then evaporated. The residue was purified by silica gel chromatography. The yield was 100 mg (100%).

For synthesis of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPPG), 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP) was dissolved in 10 mL of chloroform. To this was added 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP) and 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP). The mixture was stirred for 24 hours at room temperature. The mixture was then washed with water, dried over anhydrous calcium chloride, and then evaporated. The residue was purified by silica gel chromatography. The yield was 100 mg (100%).



Figure 1. Chemical structures of the lipids. (A) 1,2-dipalmitoyl-sn-glycero-3-phosphate (DPP), (B) 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPPG), (C) 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPPG), (D) 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPPG).

For synthesis of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPPG), 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP) was dissolved in 10 mL of chloroform. To this was added 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP) and 100 mg of 1,3-bis(sn-3'-phosphatidyl)-sn-glycero-2-phosphate (BPP). The mixture was stirred for 24 hours at room temperature. The mixture was then washed with water, dried over anhydrous calcium chloride, and then evaporated. The residue was purified by silica gel chromatography. The yield was 100 mg (100%).

MATERIALS AND METHODS

Thin Layer Chromatography. The 20- μ m-thick plates of silica gel G (Wako Pure Chemical Industries, Ltd., Richmond, BC, Canada) were used. The solvent system was chloroform:methanol:water (100:10:1) v/v/v. The plates were pre-equilibrated with the solvent system for 30 min. The samples were applied as a 10- μ l spot near the bottom. The plates were developed in the solvent system. The spots were visualized by spraying with 10% phosphomolybdic acid solution. The R_f values were recorded. The plates were dried at 100°C for 30 min.

High-Resolution Mass Spectrometry. The samples were prepared as a 10- μ l spot on a 20- μ m-thick plate of silica gel G. The plate was developed in the solvent system. The spots were visualized by spraying with 10% phosphomolybdic acid solution. The R_f values were recorded. The plates were dried at 100°C for 30 min.

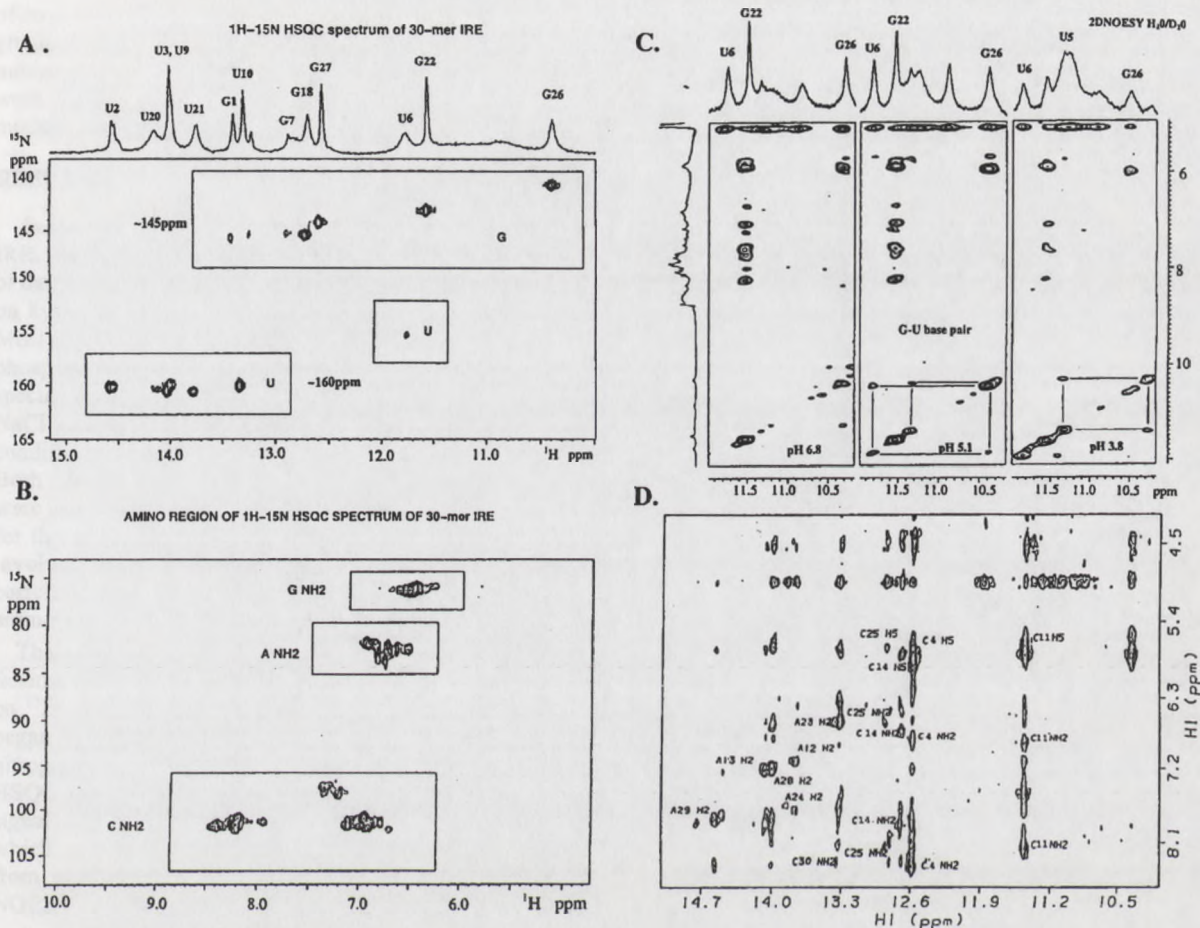


FIGURE 2: NMR data for the ferritin IRE: 2D (^1H - ^{15}N) HSQC spectra of (A) the imino region, (B) the amino region for 2D NOESY in 90% $\text{H}_2\text{O}/10\%$ D_2O , (C) imino-imino region for three different values of pH showing enhanced detection of the G-U base pair at lower pH, and (D) the imino to aromatic and amino region. RNA (0.5–1 mM) was measured at 12 $^\circ\text{C}$ in 10 mM sodium phosphate and 0.1 mM EDTA at pH 6.8.

to defocus and refocus the solvent magnetization as described by Mori et al. (28). Experiments were optimized separately for imino (G N1H and U N3H) and amino nitrogens (C, G, and A). Data sets with 5760 complex points in t_2 and 360 complex points in t_1 were acquired. In the 2D (^1H - ^{13}C) correlation experiment, constant time ^{13}C HMQC was used to enhance the resolution of the carbon dimension by refocusing the $^1\text{J}_{\text{C}-\text{C}}$ coupling with flip-back pulses for water elimination (29).

RNA samples with the native ferritin IRE sequence, containing the conserved sequence around the internal loop/bulge (Figure 1), formed opalescent gels during annealing at concentrations of >1 mM which irreversibly broadened the lines in the NMR spectrum even after diluting the sample, melting, and reannealing. Samples at concentrations below 1 mM used here and in previous studies (17) were clear under a variety of conditions and showed no evidence of intermolecular aggregation. Such properties, to date, have limited the types of multidimensional NMR experiments possible with the native ferritin IRE sequence.

Structure Modeling and Refinement. At the outset, the secondary structure of the RNA sequence was predicted using mfold version 9.0 from the University of Wisconsin GCG (Figure 1A). The program MC-SYM (macromolecular

conformation by symbolic generation) (30) was used to generate a family of model RNA/IRE structures. Preliminary structures were generated using a nucleotide conformational database constructed from all available nucleic acid structures determined by X-ray crystallography and NMR spectroscopy. The INSIGHT II program (version 95.0 from Biosym/Molecular Simulations Inc.) was then used to visualize and refine the coordinates generated by MC-SYM. Initial structures from MC-SYM were refined using energy minimization with a Discover 95.0 module. Structures were examined for violations with a limited number of NMR constraints and bad contacts. The script for the model is available upon request.

Docking. The Co(III) hexamine ion was built within INSIGHT II using the Builder module and the published crystal structure (31). An RNA/Co(III) hexamine complex was constructed using the Docking module of the Biosym/MSI software package. Automatic docking of a ligand into the IRE molecule was performed employing an energy-driven Affinity method using the FIXEDDOCKING command. FIXEDDOCKING uses nonbond methods in DISCOVER and does not include solvation terms. The extensible systematic forcefield (ESFF) was used in all DISCOVER 3 calculations. In a DOCKING search, the maximum number

of structures was set to 10. Usually, less than 400 conjugate gradient minimization steps were needed to find the energy minimum for each acceptable structure. Hydrogen bonds were located with HBOND in MEASURE in the Viewer module of INSIGHT II.

RESULTS

Exchangeable Proton Spectra of the Ferritin IRE. In the IRE, the imino (10–15 ppm) and amino (6–9 ppm) regions of the ^1H NMR spectra recorded in H_2O provided information on hydrogen bonding between base pairs and mismatches. Most of the data were recorded at pH 6.8 in 10 mM sodium phosphate buffer at 12 °C after a survey of the imino proton spectra at different temperatures, pH values (6.8–3.2), and NaCl concentrations (0–150 mM) in order to find optimal conditions for making exchangeable proton assignments. Both 2D NOESY and 2D (^1H – ^{15}N) correlated experiments were used to assign imino protons. Complete connectivities for the stem imino protons were obtained. Amino protons involved in base pairing were assigned from imino–amino correlations and ^1H – ^{15}N HSQC spectrum optimized for the amino region.

The assignment of exchangeable proton resonances in the ferritin IRE was facilitated by 2D heteronuclear experiments on ^{15}N - and ^{13}C -labeled molecules. Resonance assignment began by identifying the directly bonded imino protons and nitrogens using the imino proton optimized 2D (^1H – ^{15}N) HSQC spectrum shown in Figure 2. All but one uridine ^{15}N signal resonated in the region typical for A·U base pairs which was further confirmed by strong NOE interactions from uridine imino to adenosine H2 protons in the 2D NOESY spectrum.

Amino proton resonances for the IRE were assigned next using the amino proton optimized 2D (^1H – ^{15}N) HSQC and 2D NOESY spectra. In a G·C base pair, strong intra-base pair NOEs generally arise between the G imino protons and the C amino protons. Figure 2B shows a plot of the region containing the cross-peaks for the cytidine amino groups. Separate resonances were observed for the two protons in all but one cytidine amino group. A broad cross-peak was observed for guanosine amino proton resonances in a separate region of this spectrum, due to chemical exchange-induced line broadening caused by the rotation of the amino group around the C–N bond (19). Adenine amino groups were not well resolved in the (^1H – ^{15}N) HSQC spectrum because of an intermediate exchange regime.

Five cytidine amino groups showed characteristic bonded/nonbonded amino proton patterns indicative of involvement in Watson–Crick hydrogen bonds (Figure 2B,D). Three are in the upper and lower stems of the IRE (Figure 1). The only other two G·C base pairs possible from the secondary structure (Figure 1A) are across the hairpin hexaloop (G18·C14) and across the bulge loop region (G7·C25) below unpaired C8 and above U5/U6. The residue with the imino proton at 12.86 ppm had a correlation to C amino protons at 7.89 and 6.71 ppm, while the imino proton at 12.69 ppm showed NOEs to amino protons at 8.11 and 6.63 ppm, consistent with the restricted rotation of a C residue in a G·C base pair (32), and thus are assigned to Watson–Crick G·C base pairs. One of the imino resonances, 12.69 ppm, is assigned to a G·C pair across the hairpin hexaloop

(G18·C14) by comparison of the NMR spectra of the wild type IRE sequence and a base substitution, G18A (17). Note that the assignment of the imino proton of G18 in the hairpin loop of the IRE 30-mer differs from that of 16-mer fragment of the IRE (18) in terms of chemical shift values as well as in terms of stability, judged by the broadness of the imino signal; the difference is likely due to the influence of the internal loop/bulge and the lower stem on the IRE structure. That the structure of one region of the IRE can be sensitive to structure in other parts of the IRE is illustrated by the marked changes observed throughout the IRE in the G18A substitution (17) and by changing the concentration of magnesium [Mg^{2+}] (9). Finally, then, the fifth G·C base pair in the ferritin IRE, with an imino resonance at 12.86 ppm, is G7 paired with C25 in the internal loop/bulge region (Figure 1B). The two G·C base pairs, G7·C25 and C14·G18, each of which spans an IRE loop were not predicted by the mfold program (Figure 1A).

G26, U5, and U6 are in the internal loop/bulge region of the IRE (Figure 1). Imino proton resonances at 10.37 and 11.87 ppm could not be unambiguously assigned to either G26 or U5 or U6 from the proton spectrum alone. However, in the ^{13}C - and ^{15}N -enriched sample, the U and G imino nitrogens have distinctive chemical shifts and are readily identified by the nitrogen chemical shift. A G·U base pair can be identified by the ^{15}N chemical shifts (140.47 ppm for G and 155.3 ppm for U) in Figure 2A; note that the signal at 155.3 ppm has a very low intensity and could be observed only at the low contour plot level. The presence of the G·U base pair deduced from the ^{15}N chemical shifts is supported by the observation of a cross-peak characteristic of a G·U base pair in a 2D NOESY spectrum recorded in $\text{H}_2\text{O}/\text{D}_2\text{O}$ at 12 °C, varying the value of the pH (Figure 2C). The G·U cross-peak intensity was very sensitive to pH (Figure 2C), becoming stronger as the pH decreased from 6.8 to 5.1. Two cross-peaks characteristic of G·U base pairs were observed in the 2D NOESY spectrum (Figure 2C), indicating the occurrence of two conformers, one with a G26·U5 pair and one with a G26·U6 pair. The distribution of conformers varied with pH (Figure 2C) with the G26·U6 pair predominating at higher values of pH. The properties of the G·U base pair shown in Figure 2 suggest that the IRE in the region of the bulge/loop region of the G·U base pair is dynamic, a conclusion supported by enhanced solvent exchange (17) susceptibility to radical cleavage [Fe-EDTA/hydrogen peroxide and $\text{Cu}(\text{phen})_2/\text{dioxigen}$] (9, 16).

Model of the Ferritin IRE. MC-SYM version 1.3 was used to generate a folded starting structure of the ferritin IRE using the A type helix for the stem nucleotides G1–C4, G27–C30, U9–A13, and U20–A24, an assumption supported by the NMR data shown in Figure 2. Watson–Crick base-paired helical nucleotides were assigned a single conformation, that of a nucleotide in an idealized A-helix. Multiple conformations were allowed for the internal loop/bulge region and hexaloop nucleotides. Satisfactory distance constraints for modeling the A-helical part of the IRE were obtained from 2D NOESY experiments in H_2O (Figure 2D) and D_2O . Sugar puckers were based on data from DQF/COSY experiments in D_2O .

The loop regions of the IRE had fewer constraints. On the basis of the NMR data, C14·G18 and G7·C25 were defined as Watson–Crick base pairs; MC-SYM sampled 10

different conformations of Watson-Crick base pairs. To facilitate study of the internal loop/bulge, a hexaloop structure, CAGUGU, was generated with a script that, in addition to the C14-G18 pairs, forced G16 to stack over A15, in analogy to the hexaloop described for the IRE 16-mer subdomain (upper stem/hairpin hexaloop) studied by Liang and Hall (18). In the internal loop/bulge, C8, G7, U6, U5, and C25, a Watson-Crick base pair was included for G7-C25 in the script, on the basis of the NMR. G26-U6 was included as a base pair in the script, on the basis of the NMR data in Figure 2. Because the position of C8 could not be fixed by NMR constraints, different terms permitting a large degree of sampling were first used for this nucleotide. Structures with C8 bulged out and stacked into a helix were generated and then tested. Residue C8 is thought to be critical for recognition by the IRP since it cannot be mutated without severe loss of activity (5-7). Elimination of the bulge U6 also diminished IRP recognition (Y. Ke and E. C. Theil, unpublished observations).

The script was constructed in two stages in order to reduce the number of solutions generated by the MC-SYM script when using a large degree of sampling. First, only a 16-mer including the hexaloop (but excluding the internal loop/bulge and lower stem) was generated. Because of the limited input data, only a small number of nucleotide conformations were sampled to decrease the number of allowed solutions. Eight different hexaloop structures were generated which differed mainly in the position of U19; this position is not conserved in IREs. Cytidine was used in the 16-mer model (18) and bulged out, as does U19 in all the hexaloop models generated in this study. U17 was also looped out. The positions selected for A15, G16, and U17 are unlikely to affect greatly any other part of the IRE because the AGU sequence and position 19 are sites in the IRE hexaloop that project into the solution on the basis of the accessibility to both large and small structure probes in solution in both ferritin and transferrin receptor IREs (9, 10, 16, 38, 42).

Next, nucleotides in the hexaloop were fixed and the internal loop was modeled. The input file consisted of explicit base pairing and stacking terms for generating an A type helix (supported by the NMR data) for the stem nucleotides. For the other residues, information on stacking, non-Watson-Crick base pairs, and other conformational features served to restrict the conformational assignments to those nucleotides in the database that possess the same conformational features. In those cases where no conformational information was available, different sample conformations from the database were tested, which were compatible with known geometries of RNA structures determined by X-ray crystallography and NMR spectroscopy. The MC-SYM structures were converted into pdb format and subjected to 100 cycles of energy minimization with steepest descent and 100 cycles of minimization with conjugate gradients.

The MC-SYM script with the hexaloop fixed as described, generated 47 structures which varied in the internal loop/bulge region. The structures were grouped into five families consisting of 25, 8, 4, 5, and 5 structures. Superposition, after energy minimization, was performed for five structures representing each of the MC-SYM families (Figure 3A) and for 15 structures from the largest MC-SYM family (Figure 3B). Only the nucleotides from U3 to U20 and from A23

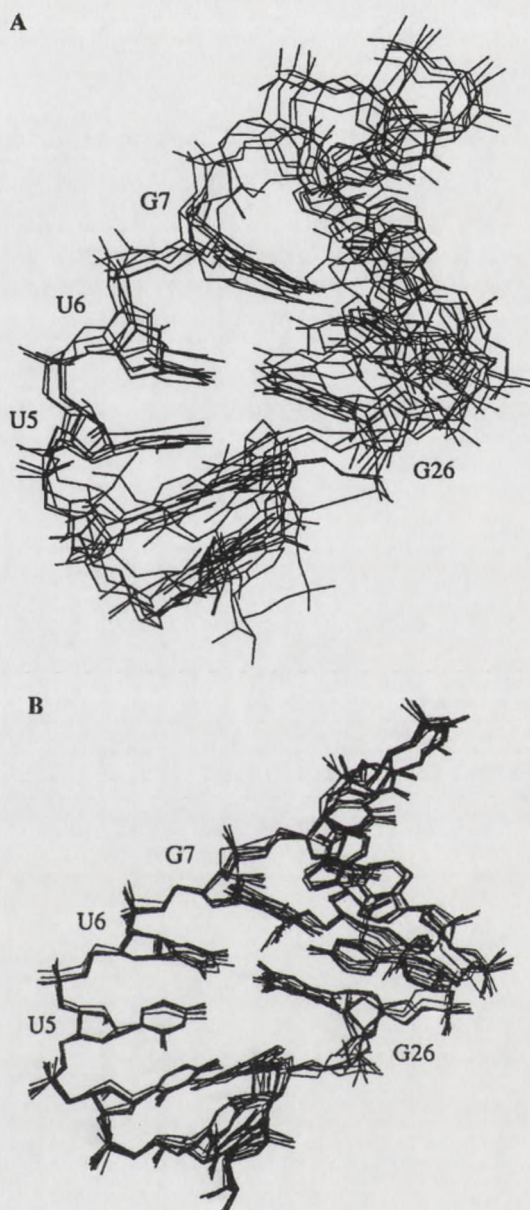


FIGURE 3: View of the internal loop/bulge region of the IRE structure. Superposition was carried out with nucleotides U3-U20 and U23-U28, which bracket the internal bulge/loop region. The view is into the minor groove: (A) five structures, one each representing the five MC-SYM families; and (B) 15 of the 25 structures in the largest MC-SYM family.

to A28, which bracket the internal loop/bulge, were used to create the best fit. All the conformations within one family are very similar, being mainly variants of position C8 and sugar-phosphate backbone conformers of G26 (Figure 3B) that likely reflect the dynamic features of the G26 and the U6-G26 base pair observed in NMR spectra (Figure 2) and the fact that U5 readily stacks in the models (Figure 3).

The model (Figure 4) shows the distortion of the helix imposed by the structure of the internal loop/bulge region, C8, G7, U6, and U5 and C25, G26, G27, and G7-C25. The distance between the G18 and G7 is 22 Å in the ferritin IRE model, and the major groove appears to be widened at the junction of the upper and lower helices. In contrast, in an

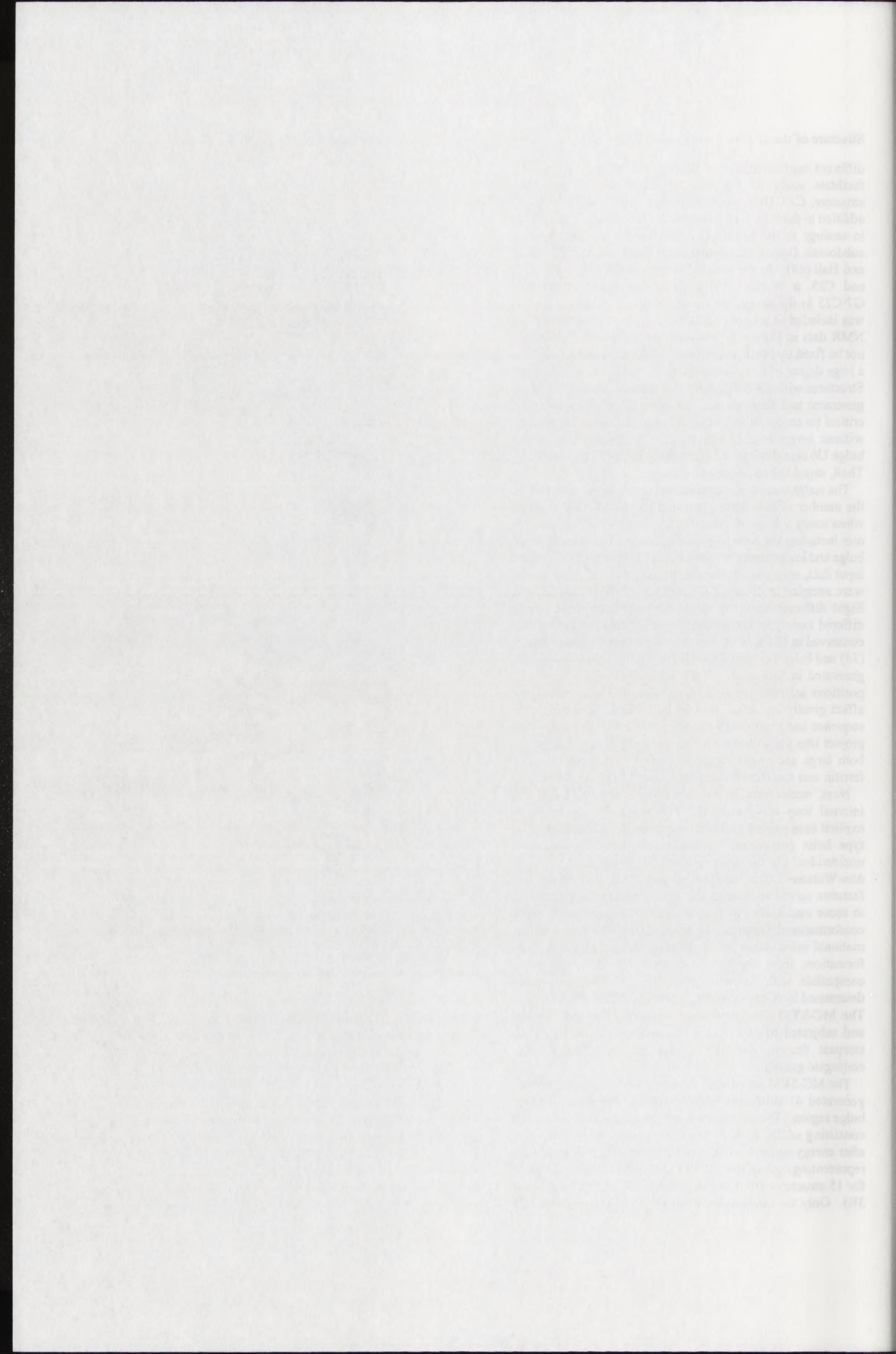




FIGURE 4: IRE models, obtained using MC-SYM, assisted by constraints obtained from NMR spectra of the IRE (Figure 2). G7 and G18, which form G-C base pairs across loops (see Figure 2), are shown in white and are at sites cross-linked to IRP-1 (21). C8, required for IRP binding, is shown in orange. U5 is stacked into the helix, while C8 appears to rotate freely (see Figure 3). The internal loop/bulge region, which is conserved in all ferritin IREs, changed the distance between the two IRP-1 contact sites from 18 to 22 Å, which may relate to IRP binding (42, 43): (left) the ferritin IRE and (right) a ferritin IRE model without the internal loop/bulge (C8 and U5 deleted).

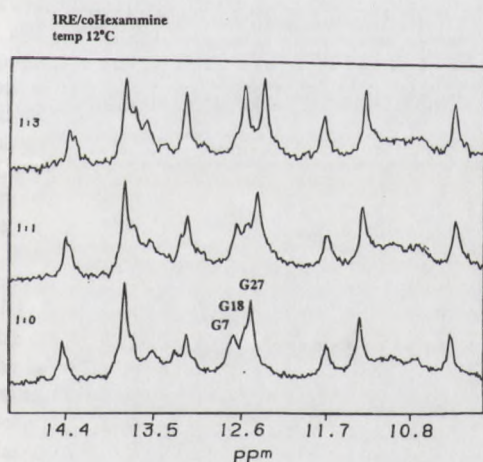


FIGURE 5: Effect of Co(III) hexammine on the ^1H NMR spectrum of the ferritin IRE. RNA (0.5 mM) was measured at 12 °C in 10 mM sodium phosphate and 0.1 mM EDTA at pH 6.8. Co(III) hexammine, present during melting and annealing, increased the T_m 4, 8, or 10 °C for 1:1, 2:1, and 3:1 Co:RNA, respectively: (top) 3:1 Co:RNA, (middle) 1:1 Co:RNA, and (bottom) 0:1 Co:RNA.

A-helix with the same sequence but without the bulges (deletion of C8 and U5), the distance between G18 into G7 is only 18 Å (Figure 4), and the major groove is narrow and undistorted. G18 and G7 are protein (IRP) contact sites, where intersite distances may be important for binding.

Co(III) Hexammine Effects on the IRE. To gain more insight about the loop regions of the IRE and to further test the model, one-dimensional NMR spectroscopy was used to monitor the Co(III) hexammine effects on the IRE. Co(III) hexammine can provide information about Mg^{2+} sites in RNA (33). Figure 5 shows the results of IRE titration with Co(III) hexammine up to a 1:3 ratio; the concentration of Co(III) hexammine was 0.5–1.5 mM. One set of imino protons of IRE and amino protons of Co(III) hexammine

(3.65 ppm) observed during the titration experiment indicated that Co(III) hexammine was in fast exchange on the NMR time scale. Addition of Co(III) hexammine to the IRE had the largest observable effects on the chemical shift of two imino resonances: G7 and G27 in the internal loop/bulge region. G27/G26 is also a major site for binding and cleavage by $\text{Cu}(\text{phen})_2$ (9, 10). Such an observation suggests that the metal ion bound specifically in the pocket formed by bases located in the internal loop/bulge region (Figure 4) near the dynamic G·U base pair. The T_m of the IRE with Co(III) hexammine was increased 4, 8, and 10 °C for 1:1, 2:1, and 3:1 metal:RNA, respectively (Y. Ke and E. C. Theil, unpublished results).

In the ferritin IRE, N7 and O6 atoms of guanine and the carbonyl O4 group of uracil residues are potential acceptor sites for hydrogen bonding with Co(III) hexammine ion. To determine how the proposed model structure of the ferritin IRE fit the NMR results of Co(III) hexammine titration, the intermolecular association between IRE and Co(III) hexammine was examined. A model of the RNA/metal ligand complex was built by automatic docking of a flexible ligand into the model structure of RNA. The representative structure (Figure 4) of the IRE was fixed in its conformation during the automatic docking process. Such an approach is consistent with recent NMR studies on the structure of the P5b stem from a group I intron ribozyme which bound the Co(III) hexammine ion and where the structures of RNA with and without Co(III) hexammine bound were found to be identical within the precision of the NMR data (33). Superposition of 10 structures of the Co(III) hexammine complex generated by the DOCKING program using an Affinity option is presented in Figure 6.

The metal ion binds to the dynamic internal loop/bulge region which forms a suitable binding pocket for Co(III) hexammine and contains four nucleotides conserved in all ferritin IREs. All structures of the generated RNA/metal complexes can be grouped into two classes. In one class, a family of hydrogen bonds, located using the HBOND option in InsightII, are found in a complicated interaction with G27, G26, U5, U6, and C25; in the second class, additional hydrogen bonds are found to G7. Positioning C8, in or outside the bulge loop, had no effect on the model. All the hydrogen bonds in the complexes are to base nitrogen and oxygen; no interactions with the backbone were observed in the model. The Co(III) hexammine:IRE ratio (1:1) increases the T_m by 4 °C (Y. Ke and E. C. Theil, unpublished observations). The effect of Co(III) hexammine on T_m , combined with the effect of Co(III) hexammine on the NMR spectrum and the model, suggests that the Co(III) hexammine interaction at the dynamic G·U base pair selects and stabilizes one of the conformations around G26 that are possible in the model (Figure 3).

DISCUSSION

A model of the natural IRE in ferritin mRNA, generated by a combination of molecular modeling (MC-SYM) and NMR studies (Figures 2 and 3), is a stack of two A-helices of 5 base pairs each, distorted by an internal loop/bulge region with a terminal hairpin loop. Earlier NMR studies of the IRE (17, 18) considered only the terminal hexaloop and obtained evidence for the G·C base pair (17, 18) also

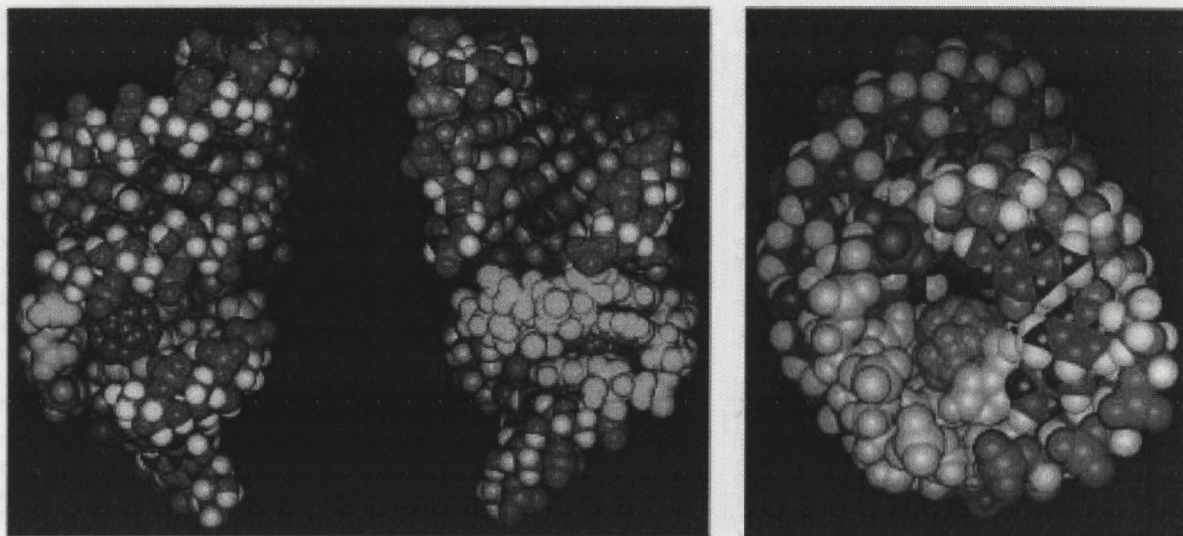


FIGURE 6: Structure of the IRE/Co(III) hexamine complex. Docking was used to examine the interaction of Co(III) hexamine (blue) with the IRE. Note the specific interactions with G7, G27, G26, U5, U6, and C25 (yellow) in the pocket of the internal loop/bulge region. ^1H resonances of G7 and G27 on the outer edge of the pocket were the most affected by Co(III) hexamine binding, based on the NMR spectra (Figure 5). G18 in the terminal hexaloop is shown in orange: (A, left structure in left panel) view into the minor groove and (B) view from the bottom (5'- and 3'-termini).

observed here (Figure 2). Solution studies of the ferritin IRE using RNAses, dimethyl sulfate, transition metals complexes that are nucleases, and SELEX of variants of the hairpin loop with an IRP-1 ligand (9, 10, 16, 34–38) are all consistent with the model. Both the internal loop/bulge and the hairpin loop are spanned by a G-C base pair (G7-C25 and C14-G18, respectively) (Figure 1B). The G-C base pairs observed in the NMR spectra (Figure 2) are between bases conserved in all ferritin IREs, but were not predicted (Figure 1A); the G-C pair in the terminal hexaloop had also been indicated in earlier chemical and mutagenesis studies (9, 10, 16–18, 34–38), but before this study there was no experimental evidence for the G-C pair in the internal loop/bulge. On the basis of the imino resonances (12.69 and 12.86 ppm), the relatively high accessibility to D_2O , and relatively low temperature stability (17), the two G-C base pairs are in similar environments. Moreover, each base pair is adjacent to a dynamic region of the IRE (U6, U5, C8, or U19) and protein contact sites (21), suggesting that a combination of structural specificity and flexibility is important for the RNA/protein interaction.

Residue C8, which does not participate in a Watson-Crick base pair and is conserved in all IREs (5, 6, 8), could not be fixed by NMR constraints. Free rotation of the amino group of cytosine residues is limited only to the unpaired form of the base and is represented by the single cytidine amino group in the 2D (^1H - ^{15}N) NMR data (Figure 2B). Thus, the flexibility of C8, and consequent free rotation of the amino group, may be related to a requirement of water-mediated protein complexation (32).

Co(III) hexamine binds to the ferritin IRE where the major groove is distorted by the internal loop/bulge of the ferritin IRE, based on the effect on the 1D NMR spectrum (Figure 5). Modeling the Co(III) hexamine/RNA interaction (Figure 6) provided a more detailed view of the internal loop/bulge region and the Co(III) hexamine site than is possible from the NMR data alone. The interaction of the

IRE with Co(III) hexamine appears to be analogous to the binding of Co(III) hexamine in the major groove of the P5b stem of the group I intron of *Tetrahymena* (33). In the ferritin IRE, the Co(III) hexamine site is located in the major groove in a pocket formed by a G-U base pair and adjacent G and U residues; in the model, hydrogen bonds between Co(III) hexamine and U5, U6, G7, G26, and G27 are predicted. $\text{Cu}(\text{phen})_2$ appears to bind in or near the Co(III) hexamine site of the ferritin IRE on the basis of cleavage of G7, G26, and G27 in both natural ferritin mRNA [poly(A⁺)] and *in vitro* transcripts (9, 35, 36). $\text{Cu}(\text{phen})_2$ recognizes distortions in RNA helices associated with bulge loops (39, 40).

Hydrated Mg^{2+} and Co(III) hexamine can bind at similar sites in RNA, often involving G residues, as illustrated by the group I intron of *Tetrahymena* (33, 41). By analogy, the Co(III) hexamine site in the internal loop/bulge of the ferritin IRE could also be a Mg^{2+} binding site. The fact that Mg^{2+} decreased the accessibility of $\text{Cu}(\text{phen})_2$ to G26 and G27 in the internal loop/bulge region (9) suggests competition between Mg^{2+} and $\text{Cu}(\text{phen})_2$ for the same IRE site and strengthens the idea that the internal loop/bulge region of the ferritin IRE is a Mg^{2+} binding site. Since the IRE forms a fairly specific structure without Mg^{2+} or Co(III) hexamine (Figures 2 and 4) at low salt concentrations, finding conditions where the effect of metal or salt on IRE function can be studied independently of effects on protein is difficult, except by mutation (38; reviewed in refs 5–8). The relatively stable structure of the IRE without metal, except for the internal loop/bulge, suggests that the metal will be involved in fine-tuning interactions in an initial RNA/protein complex, especially in the region of the internal loop/bulge, or may facilitate protein dissociation.

Hydrogen bonds between Co(III) hexamine and the IRE molecule involving five of the six NH_3^+ groups and U5, U6, G26, and G27 are possible. Such a hydrogen bond network could stabilize the dynamic features of base pairing exempli-



Figure 1. A line graph showing the relationship between [illegible] and [illegible]. The x-axis is labeled [illegible] and the y-axis is labeled [illegible]. The data points show a positive correlation, with a regression line drawn through them.

The first part of the study was designed to determine the effect of [illegible] on [illegible]. The results showed that [illegible] had a significant effect on [illegible]. The second part of the study was designed to determine the effect of [illegible] on [illegible]. The results showed that [illegible] had a significant effect on [illegible]. The third part of the study was designed to determine the effect of [illegible] on [illegible]. The results showed that [illegible] had a significant effect on [illegible].

The results of this study indicate that [illegible] has a significant effect on [illegible]. This finding is important because [illegible]. The study was limited by [illegible]. Further research is needed to [illegible].

fied in Figure 2 and could account for the increase in the T_m when Co(III) hexammine binds (Y. Ke and E. C. Theil, unpublished observations). The interhelical region of IREs appears to be different in other IREs, based on the predicted structure which has only the C8 bulge. However, the IREs in erythroid aminolevulinic acid and *m*-aconitase mRNAs have additional bulged residues near the C bulge which may produce that same dynamic behavior of the interhelical region observed with the ferritin IRE (Figure 2C). The transferrin receptor IREs, which have a bulge C and no other predicted bulge regions, nevertheless have a broad melting transition, multiple cleavage sites [Cu(phen)₂] (Ke and E. C. Theil, unpublished observations), and multiple Pb²⁺ hydrolysis sites (42) in the IRE interhelical region, which also is the behavior of an RNA with conformational heterogeneity in the interhelical region. Taken together, the NMR data, the IRE model, and the IRE chemical reactivity suggest that a fairly commodious or dynamic metal binding pocket in the middle of the IRE helix is important for optimal ligand/protein interactions.

NOTE ADDED IN PROOF

A paper describing the structure of a consensus B(bulge)-IRE appeared recently (43). Many of the features are shared except in the region of the IL/B (internal loop/bulge); regulation by the IL/B-IREs in vivo is more efficient than that by B-IREs.

REFERENCES

- Peterson, R. D., Bartel, D. P., Szostak, J. W., Horvath, S. J., and Feigon, J. (1994) *Biochemistry* 33, 5357–5366.
- Puglisi, J. D., Chen, L., Blanchard, S., and Frankel, A. D. (1995) *Science* 270, 1200–1203.
- Brodsky, A. S., and Williamson, J. R. (1997) *J. Mol. Biol.* 267, 624–639.
- Luebke, K. J., and Tinoco, I., Jr. (1996) *Biochemistry* 35, 11677–11684.
- Roault, T. A., and Klausner, R. D. (1996) *J. Biol. Inorg. Chem.* 1, 494–499.
- Hentze, M. W., and Kuhn, L. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 8175–8182.
- Goessling, L. S., Rup, D., and Thach, R. E. (1996) *Prog. Nucleic Acids Res. Mol. Biol.* 55, 121–134.
- Theil, E. C. (1997) in *Metal Ions in Biological Systems* (Sigel, H., and Sigel, A., Eds.) (in press).
- Wang, Y.-H., Sczekan, S. R., and Theil, E. C. (1990) *Nucleic Acids Res.* 18, 4463–4468.
- Wang, Y.-H., Lin, P.-N., Sczekan, S. R., McKenzie, A. R., and Theil, E. C. (1991) *Biol. Met.* 4, 56–61.
- Shull, G. E., and Theil, E. C. (1982) *J. Biol. Chem.* 257, 14187–14191.
- Dix, D. J., Lin, P.-N., Kimata, Y., and Theil, E. C. (1992) *Biochemistry* 31, 2818–2822.
- Bhasker, C. R., Burgiel, G., Neupert, B., Emery-Goodman, A., Kuhn, L. C., and May, B. K. (1993) *J. Biol. Chem.* 268, 12699–12705.
- Shull, G. E., and Theil, E. C. (1983) *J. Biol. Chem.* 258, 7921–7923.
- Dickey, L. F., Wang, Y.-H., Shull, G. E., Wortman, I. A., III, and Theil, E. C. (1988) *J. Biol. Chem.* 263, 3071–3074.
- Harrell, C. M., McKenzie, A. R., Patino, M. M., Walden, W. E., and Theil, E. C. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 4166–4170.
- Sierzputowska-Gracz, H., McKenzie, A. R., and Theil, E. C. (1995) *Nucleic Acids Res.* 23, 146–153.
- Liang, L. G., and Hall, K. B. (1996) *Biochemistry* 35, 13586–13596.
- Varani, G., Aboul-ela, F., and Allain, H.-T. (1996) *Prog. Nucl. Magn. Reson. Spectrosc.* 29, 51–127.
- Pardi, A. (1995) *Methods Enzymol.* 261, 350–386.
- Basilion, J. P., Rouault, T. A., Massinople, M. C., Klausner, R. D., and Burgess, H. W. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 574–578.
- Studier, F. W., Rosenberg, A. H., Dunn, J. J., and DuBendary, J. W. (1990) *Methods Enzymol.* 185, 60–88.
- Hayne, S. L., and Whitesides, G. M. (1990) *Appl. Biochem. Biotechnol.* 23, 205–222.
- Nikonowicz, E. P., Sirt, A., Legault, P., Jackson, F. M., Baer, L. M., and Pardi, A. (1992) *Nucleic Acids Res.* 20, 4507–4513.
- Batay, R. T., Foruda, M., Kujawinski, E., Puglisi, J. D., and Williams, J. (1992) *Nucleic Acids Res.* 20, 4515–4523.
- Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* 2, 661–665.
- Sklenar, V., and Bax, A. (1987) *J. Magn. Reson.* 74, 469–479.
- Mori, S., Abeygunawardana, C., Johnson, M. O., and van Zijl, P. C. M. (1995) *J. Magn. Reson., Ser. B* 108, 94–98.
- Hallenga, K., and Lippens, G. (1995) *J. Biomol. NMR* 5, 59–66.
- Major, F., Turcotte, M., Gautheret, D., Lapalme, G., Fillion, E., and Cedergren, R. (1991) *Science* 253, 1255–1260.
- Clegg, W. (1985) *Acta Crystallogr. C* 41, 1164–1166.
- Williams, L. D., Williams, N. G., and Shaw, B. R. (1990) *J. Am. Chem. Soc.* 112, 829–832.
- Kieft, J. J., and Tinoco, I., Jr. (1997) *Structure* 5, 713–721.
- Bettany, A. J. E., Eisenstein, R. S., and Munro, H. M. (1992) *J. Biol. Chem.* 267, 16531–16537.
- Dix, D. J., Lin, P.-N., McKenzie, A. R., Walden, W. E., and Theil, E. C. (1993) *J. Mol. Biol.* 231, 230–240.
- Henderson, B. R., Menotti, E., Bonnard, C., and Kuhn, L. C. (1994) *J. Biol. Chem.* 269, 17481–17489.
- Butt, J., Kim, H.-Y., Basilion, J. P., Cohen, S., Iwai, K., Philpott, C. C., Altschul, S., Klausner, R. D., and Rouault, T. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 4345–4349.
- Thorp, H. H., McKenzie, R. A., Lin, P.-N., Walden, W. E., and Theil, E. C. (1996) *Inorg. Chem.* 35, 2773–2779.
- Sigman, D. S. (1990) *Biochemistry* 29, 9097–9105.
- Hermann, T. H., and Heumann, H. (1995) *RNA* 1, 1009–1017.
- Cate, J. H., Gooding, A. R., Powell, E., Zhou, K., Golden, B. L., Kuindrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science* 273, 1678–1686.
- Schlegel, J., Gegout, V., Schlager, B., Hentze, M. W., Westhof, E., Ehresmann, C., Ehresmann, B., and Romby, P. (1997) *RNA* 3, 1159–1172.
- Address, K. J., Basilion, J. P., Klausner, R. D., Rouault, T. A., and Pardi, A. (1997) *J. Mol. Biol.* 274, 72–83. BI9719814

BI9719814

The first part of the paper is devoted to a review of the literature on the subject. It is found that the majority of the studies have been concerned with the effects of the various factors on the rate of the reaction. The results of these studies are summarized in Table I. It is seen that the rate of the reaction is affected by the concentration of the reactants, the temperature, and the presence of a catalyst. The effect of the concentration of the reactants is in accordance with the law of mass action. The effect of the temperature is in accordance with the Arrhenius equation. The effect of the presence of a catalyst is to increase the rate of the reaction.

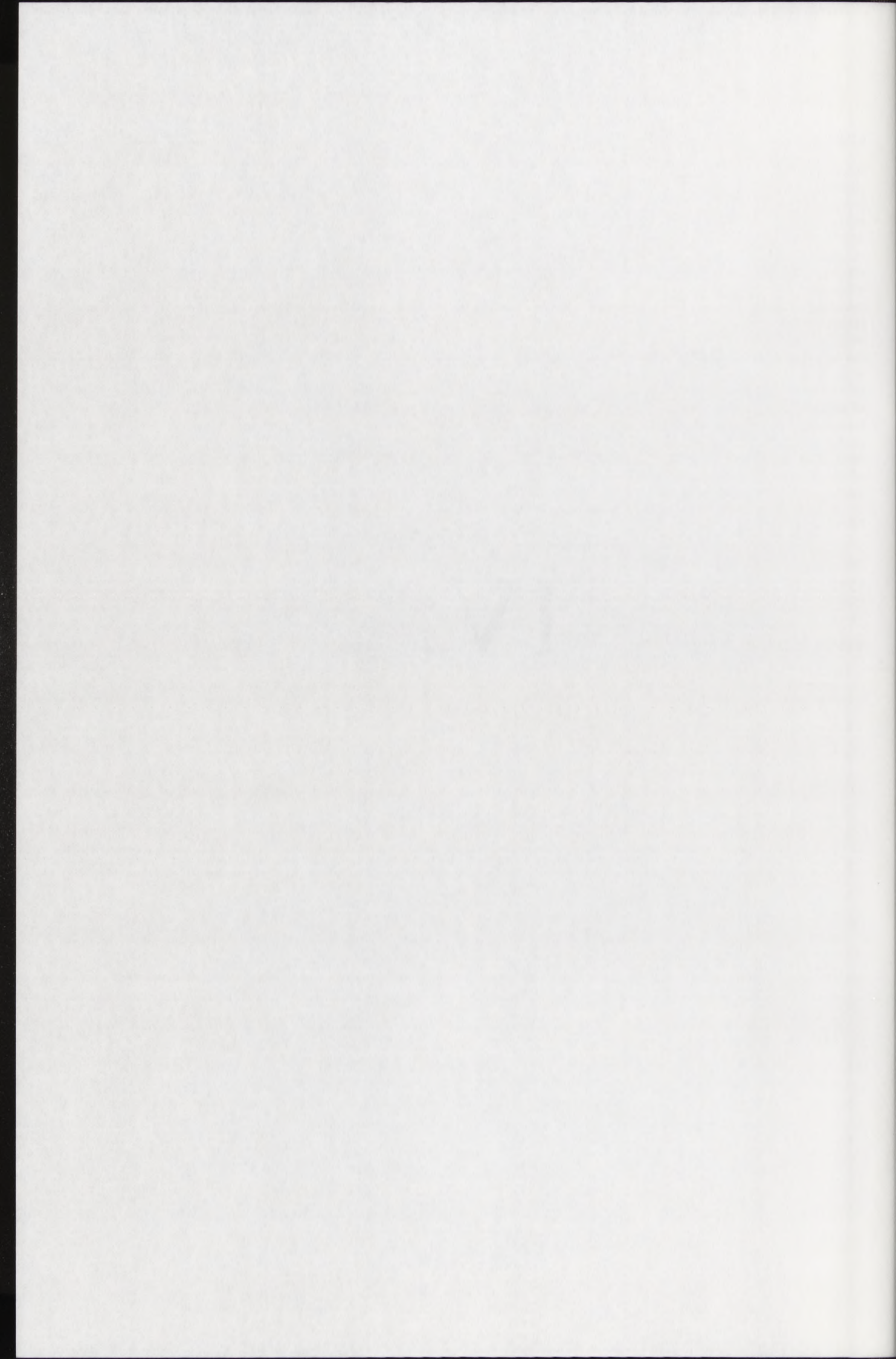
NOTE ADDED IN PROOF

A note added in proof has been received from the author, which states that the results of the present study are in agreement with those of other workers in the field.

REFERENCES

1. Smith, J. D., *J. Chem. Phys.*, **10**, 442 (1942).
2. Jones, K. L., *J. Chem. Phys.*, **11**, 100 (1943).
3. Brown, A. G., *J. Chem. Phys.*, **12**, 100 (1944).
4. Taylor, H. S., *J. Chem. Phys.*, **13**, 100 (1945).
5. Hill, J. A., *J. Chem. Phys.*, **14**, 100 (1946).
6. Fisher, J. W., *J. Chem. Phys.*, **15**, 100 (1947).
7. Goulet, J. P., *J. Chem. Phys.*, **16**, 100 (1948).
8. Hill, J. A., *J. Chem. Phys.*, **17**, 100 (1949).
9. Taylor, H. S., *J. Chem. Phys.*, **18**, 100 (1950).
10. Fisher, J. W., *J. Chem. Phys.*, **19**, 100 (1951).
11. Goulet, J. P., *J. Chem. Phys.*, **20**, 100 (1952).
12. Hill, J. A., *J. Chem. Phys.*, **21**, 100 (1953).
13. Taylor, H. S., *J. Chem. Phys.*, **22**, 100 (1954).
14. Fisher, J. W., *J. Chem. Phys.*, **23**, 100 (1955).
15. Goulet, J. P., *J. Chem. Phys.*, **24**, 100 (1956).

IV.



Internal Loop/Bulge and Hairpin Loop of the Iron-Responsive Element of Ferritin mRNA Contribute to Maximal Iron Regulatory Protein 2 Binding and Translational Regulation in the Iso-iron-responsive Element/Iso-iron Regulatory Protein Family

Yaohuang Ke,[†] Hanna Sierzputowska-Gracz,[§] Zofia Gdaniec,^{||} and Elizabeth C. Theil^{*‡}

CHORI (Children's Hospital Oakland Research Institute), Oakland, California 94609-1673, Department of Biochemistry and Chemistry, North Carolina State University, Raleigh, North Carolina 27696-7622, and Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland

Received October 25, 1999; Revised Manuscript Received February 16, 2000

ABSTRACT: Iron-responsive elements (IREs), a natural group of mRNA-specific sequences, bind iron regulatory proteins (IRPs) differentially and fold into hairpins [with a hexaloop (HL) CAGUGX] with helical distortions: an internal loop/bulge (IL/B) (UGC/C) or C-bulge. C-bulge iso-IREs bind IRP2 more poorly, as oligomers ($n = 28-30$), and have a weaker signal response in vivo. Two trans-loop GC base pairs occur in the ferritin IRE (IL/B and HL) but only one in C-bulge iso-IREs (HL); metal ions and protons perturb the IL/B [Gdaniec et al. (1998) *Biochemistry* 37, 1505–1512]. IRE function (translation) and physical properties (T_m and accessibility to nucleases) are now compared for IL/B and C-bulge IREs and for HL mutants. Conversion of the IL/B into a C-bulge by a single deletion in the IL/B or by substituting the HL CG base pair with UA both derepressed ferritin synthesis 4-fold in rabbit reticulocyte lysates (IRP1 + IRP2), confirming differences in IRP2 binding observed for the oligomers. Since the engineered C-bulge IRE was more helical near the IL/B [Cu(phen)₂ resistant] and more stable (T_m increased) and the HL mutant was less helical near the IL/B (ribonuclease T1 sensitive) and less stable (T_m decreased), both CG trans-loop base pairs contribute to maximum IRP2 binding and translational regulation. The ¹H NMR spectrum of the Mg-IRE complex revealed, in contrast to the localized IL/B effects of Co(III) hexaammine observed previously, perturbation of the IL/B plus HL and interloop helix. The lower stability and greater helix distortion in the ferritin IL/B-IRE compared to the C-bulge iso-IREs create a combinatorial set of RNA/protein interactions that control protein synthesis rates with a range of signal sensitivities.

Awareness of mRNA regulation as a mechanism for controlling gene expression is increasing. The iso-IRE (iron responsive element)¹ family of mRNA regulatory elements recognized by the cognate proteins, iron regulatory proteins (IRPs), is one of the most extensively characterized mRNA regulatory targets. However, knowledge about mRNA regulation is in its infancy compared to understanding of DNA regulation such as hormone response elements recognized by hormone nuclear receptors (1). The ancient nature of the proteins encoded in IRE-containing, animal mRNAs (e.g., aconitase, ferritin), the homology of the IRPs to aconitases, and the recent detection of functional IREs and IRPs in bacteria (2) as well as animals suggest that the IRE/IRP interaction is also ancient, possibly representing regulation in an RNA world.

The common structural features of iso-IREs are a hairpin hexaloop (CAGUGX) with a trans-loop GC base pair and a helical stem (3–5). The GC base pair in the hexaloop has a large effect on the stabilization of the overall IRE structure (3, 6). A disordered C residue occurs in all IREs (4, 5). In many iso-IREs, the disordered C is a bulge in the helix of the IRE stem (4). In another type of iso-IRE, the disordered C residue is part of a set of four conserved residues [UGC-(16 nucleotides)-C] that form an internal loop/bulge (IL/B) with a trans-loop GC base pair (5). IRE sequences are mRNA-specific. IRE sequence differences in a single organism range from 36% to 85%, but for a particular mRNA, the IRE has high sequence conservation among animals, especially among vertebrates (>95%) (7).

Common structural features of iso-IRPs include protein kinase sites (8, 9) and sequence homology to aconitases (10), although no three-dimensional IRP structures are determined at this time. IRP1 binds Fe and sulfur with the acquisition of aconitase activity as cytosolic aconitase (10).

IRE-containing mRNAs have quantitatively different responses in vivo to an environmental signal such as iron (11, 12), allowing a range of response sensitivities to the same signal. Recently, in vitro studies of the IRE/IRP interaction, using short RNA sequences (28–30 nucleotides) from several different IRE-containing mRNAs and both recombinant and natural IRP1 and IRP2, showed different

[†] Partial support for this work came from NIH R01-DK-20251 and the CHORI Foundation.

* Corresponding author: CHORI (Children's Hospital Oakland Research Institute), 5700 Martin Luther King, Jr. Way, Oakland, CA 94609-1673. Tel (510) 450-7670; Fax (510) 597-7131; E-mail etheil@chori.org.

[‡] CHORI.

[§] NC State University.

^{||} Polish Academy of Sciences.

¹ Abbreviations: IRE, iron-responsive element; IRP, iron regulatory protein; IL/B, internal loop/bulge; HL, iron-responsive element hairpin hexaloop (Hairpin Loop); RRL, rabbit reticulocyte lysate; TfR, transferrin receptor; WGE, wheat germ extract; WT, wild type.

Internal Looping of the H₂O₂ mRNA Controls Its Translation Regulation in the Yeast *S. cerevisiae*

Yoshiko Kato and Toshiyuki Nishida
Department of Applied Chemistry, Faculty of Engineering, Kansai University, Suita, Osaka 564-8680, Japan

Received 12 October 1998; Accepted 15 December 1998

ABSTRACT The yeast *S. cerevisiae* H₂O₂ mRNA contains an internal loop structure in the 5' UTR, which is essential for the regulation of its translation. The internal loop structure is located in the 5' UTR of the H₂O₂ mRNA, and its formation is dependent on the presence of a specific sequence in the 5' UTR. The internal loop structure is essential for the regulation of the translation of the H₂O₂ mRNA, and its formation is dependent on the presence of a specific sequence in the 5' UTR. The internal loop structure is essential for the regulation of the translation of the H₂O₂ mRNA, and its formation is dependent on the presence of a specific sequence in the 5' UTR.

Introduction The yeast *S. cerevisiae* H₂O₂ mRNA contains an internal loop structure in the 5' UTR, which is essential for the regulation of its translation. The internal loop structure is located in the 5' UTR of the H₂O₂ mRNA, and its formation is dependent on the presence of a specific sequence in the 5' UTR. The internal loop structure is essential for the regulation of the translation of the H₂O₂ mRNA, and its formation is dependent on the presence of a specific sequence in the 5' UTR.

Materials and Methods The yeast *S. cerevisiae* H₂O₂ mRNA was purified from total RNA by poly(A)-dependent methods. The internal loop structure was analyzed by Northern blotting and RNA electrophoresis. The translation of the H₂O₂ mRNA was measured by immunoprecipitation and Western blotting.

Results and Discussion The internal loop structure is essential for the regulation of the translation of the H₂O₂ mRNA, and its formation is dependent on the presence of a specific sequence in the 5' UTR.

Conclusion The internal loop structure is essential for the regulation of the translation of the H₂O₂ mRNA, and its formation is dependent on the presence of a specific sequence in the 5' UTR.

Keywords H₂O₂ mRNA, internal loop, translation regulation, yeast *S. cerevisiae*.

IRP2 binding among iso-IREs: the RNA/protein complexes and unbound RNA were separated by electrophoresis in native, sieving gels (13). The iso-IRE/IRP binding differences coincided with the range of different responses to iron in vivo (11–13). The differences in IRP2 binding to iso-IREs also correlated with differences between the IL/B and C-bulge and were abrogated by deletion of a conserved U residue from the internal loop/bulge to create a C-bulge (13).

To begin to connect the data on IRP binding for the short RNA sequence in vitro with the observed range of iron responsiveness in vivo, we now show that IRE-dependent regulation of protein synthesis in vitro reproduced the differences previously observed for iso-IRE binding to IRP1 and IRP2. In addition, decreases in translational repression and IRP2 binding were associated with either more or less distortion in the helix and more or less stability of the IRE structure, on the basis of $\text{Cu}(\text{phen})_2$ or T1 probing and T_m measurements. Finally, NMR spectroscopy again emphasized the local sensitivity of the IL/B to metal ions observed previously (5) but also, since the hexaloop and interloop helix residues were perturbed, revealed the more global effects of Mg.

MATERIALS AND METHODS

Mutagenesis. IRE mutants in full-length transcripts, used to study IRE function in regulation of protein synthesis, were generated on the pBFH-1DV plasmid, which had been constructed by insertion of wild-type bullfrog H-subunit cDNA into pTZ19U vector (14, 15), by double-stranded, site-directed mutagenesis. The mutagenesis was done with a Chameleon double-stranded, site-directed mutagenesis kit (Stratagene) and followed the protocol provided by the manufacturer. The selection primer (5'-CTG TGA CTG GTG ACG CGT CAA CCA AGT C-3') was designed to change the restriction endonuclease *ScaI* recognition site to the *MluI* site for enrichment and selection of IRE mutants. There is only one *ScaI* site in the vector of pBFH-1DV plasmid. The mutagenic primers were designed to produce IRE mutants. The sequences used are 5'-CAC TGT AGC AGA ACT CTA CTA AGA G-3' (ΔU6 mutant), 5'-GGG TTC CGT TCA AAT ACT ATT GAA GCA AGA ACT CTA C-3' (C14U/G18A mutant), and GGG TTC CGT TCA AAC ACT ATT GAA GCA AGA ACT CTA C (G18A mutant). The IRE mutants were confirmed by DNA sequencing with a T7 Sequenase version 2.0 DNA sequencing Kit (U.S. Biochemical Corp.).

RNA Preparation. Capped full-length RNA transcripts were transcribed from pBFH-1DV and the derived mutant (pBFH- ΔU6 , pBFH-C14U/G18A, and pBFH-G18A) generated by mutagenesis. The plasmids were digested with *SaII* restriction endonuclease, 24 nucleotides downstream from the DNA sequence encoding the ferritin mRNA poly(A)+ tail. The linearized plasmids were purified by extraction with phenol, phenol/chloroform, and chloroform, precipitated with ammonium acetate and ethanol, and then used as templates for transcription. Transcription reactions were performed as described previously (14–16) except for the omission of radioactive nucleotide in the reactions. The transcription products were purified through RNeasy columns (Qiagen) and eluted in H_2O . The purified transcript was quantitated by reading the absorbance at 260 nm at various dilutions

and stored at -80°C . Homogeneity of the transcripts was analyzed on 1 M urea/1.5% agarose gels.

RNAs (29–30-mer) used for thermal denaturation and enzymatic and chemical structure probing were transcribed from chemically synthesized DNA templates with T7 RNA polymerase (6, 17), followed by gel purification and ethanol precipitation, and then resuspended in H_2O and stored at -80°C until use. RNA labeling was performed as described previously (18, 19), with purification through NEN sorb columns (DuPont). The RNAs used for thermal denaturation were further purified by dialysis against H_2O .

For NMR analysis, the ferritin IRE 30-mer was synthesized and gel-purified by the CyberSyn Company in Lenni, PA. Samples for the analysis were prepared by dissolving approximately 1.1 mg of the purified RNA in 0.2 mL of 90% $\text{H}_2\text{O}/10\%$ D_2O phosphate buffer (10 mM sodium phosphate, pH 6.8, and 0.1 mM EDTA) to give ~ 0.5 mM RNA. The samples were heated at 80°C for 5 min and cooled slowly prior to each experiment. Mg^{2+} was added after annealing. Aliquots of the stock solution of 0.3 mM MgCl_2 were individually added to the RNA sample to produce different ratios of the divalent ions to RNA. At the strand concentration used $\gg 90\%$ of the RNA was monomeric, on the basis of melting profiles and diffusion NMR analysis.

Protein Synthesis in Vitro. Nuclease-treated rabbit reticulocyte lysates (RRL) and wheat germ extracts (WGE) from Promega were used for protein synthesis as previously described (14, 15, 20, 21). Ferritin synthesis with [^{35}S]-methionine labeling in RRL was carried out at 30°C for 17 min and in WGE at 25°C for 40 min. The final concentration of full-length transcripts directing ferritin synthesis is 1.5 nM in RRL and 3.0 nM in WGE. Ferritin synthesis rates were linearly dependent on the incubation time and RNA concentration on the basis of preliminary experiments with different incubation times and RNA concentrations. After termination of the labeling translation reactions with an equal volume of 10 mM cold methionine, the labeled, synthesized ferritin was resolved by SDS-polyacrylamide gel electrophoresis and analyzed with a PhosphorImager and ImageQuant software (Molecular Dynamics).

Enzymatic and Chemical Structure Probing. Cleavage of RNA with T1 and $\text{Cu}(\text{phen})_2$ complexes followed the previous description (18, 19) with modifications. 5'- ^{32}P -Labeled RNAs were melted at 85°C for 5 min and then slowly annealed in a metal block at room temperature before incubation with RNase T1 or $\text{Cu}(\text{phen})_2$ under native conditions. Cleavage by the transition metal complex was initiated with 3-mercaptopropionic acid at 25°C and terminated with 2,9-dimethyl-1,10-phenanthroline after 5 min. Digestion with T1 under denaturing conditions and alkaline hydrolysis of RNA for gel calibration have been previously described (18). The cleavage products were resolved by electrophoresis in a denaturing gel (18) and analyzed with a PhosphorImager and ImageQuant software (Molecular Dynamics).

Thermal Denaturation. Wild-type or mutant IRE (~ 1.5 μM in 10 mM phosphate buffer, pH 6.8, with or without 100 mM NaCl) were heated and annealed, as described under Enzymatic and Chemical Structure Probing, prior to thermal denaturation analysis. RNA thermal denaturation over the range of 5 – 95°C used 1 cm path length, reduced-volume,

The first part of the study was conducted in a laboratory setting. The subjects were divided into two groups: a control group and an experimental group. The control group received a standard treatment, while the experimental group received a modified treatment. The results of the laboratory study showed that the modified treatment was significantly more effective than the standard treatment in terms of reducing the number of side effects.

In the second part of the study, the effectiveness of the modified treatment was evaluated in a clinical trial. The subjects were again divided into two groups: a control group and an experimental group. The control group received the standard treatment, and the experimental group received the modified treatment. The results of the clinical trial showed that the modified treatment was significantly more effective than the standard treatment in terms of reducing the number of side effects. The modified treatment was also found to be more cost-effective than the standard treatment.

MATERIALS AND METHODS

The study was conducted in a laboratory setting. The subjects were divided into two groups: a control group and an experimental group. The control group received a standard treatment, while the experimental group received a modified treatment. The results of the laboratory study showed that the modified treatment was significantly more effective than the standard treatment in terms of reducing the number of side effects.

The study was conducted in a clinical trial. The subjects were divided into two groups: a control group and an experimental group. The control group received the standard treatment, and the experimental group received the modified treatment. The results of the clinical trial showed that the modified treatment was significantly more effective than the standard treatment in terms of reducing the number of side effects. The modified treatment was also found to be more cost-effective than the standard treatment.

The study was conducted in a laboratory setting. The subjects were divided into two groups: a control group and an experimental group. The control group received a standard treatment, while the experimental group received a modified treatment. The results of the laboratory study showed that the modified treatment was significantly more effective than the standard treatment in terms of reducing the number of side effects.

quartz cells (800 μ L). The thermal denaturation profiles, absorbance at 260 nm as a function of temperature, were acquired on a Cary 100 spectrophotometer system with accessories for thermal melting. The rate of temperature increase was 1 $^{\circ}$ C/min. Melting temperature (T_m) and free energy (ΔG_{37}) were determined with hypochromicity methods (Cary Win Bio package application software).

NMR Spectroscopy. NMR spectra were obtained on a Bruker AVANCE 500 MHz spectrometer (1990) with an Oxford narrow-bore magnet (1989), SGI INDY Host Workstation, XWINNMR software. Spectra in 10%/90% D₂O/H₂O were acquired with a triple-gradient probe and WATERGATE pulse sequence for water suppression (22). Spectra acquired at 4 or 12 $^{\circ}$ C showed no detectable difference. Spectra were processed with XWINNMR software and Felix software (MSI), using exponential weighting function or shifted sine-bell function to resolve overlapped imino protons.

RESULTS

Effects of Mutations in the Internal Loop/Bulge or Hairpin Loop on Translation Repression. To determine if the internal loop/bulge (IL/B) and C-bulge IREs' effect on translational regulation of protein synthesis was the same as on IRP2 binding in vitro (13), ferritin synthesis directed by 5'-capped, 3'-polyadenylated, full-length ferritin mRNA was compared for RNA with wild type (IL/B) IRE or Δ U6 (C-bulge) IRE sequences (Figure 1). Previous studies had shown that the engineered C-bulge ferritin IRE, Δ U6, displays the same type of IRP2 binding as natural C-bulge IREs (erythroid aminolevulinic synthase, mitochondrial aconitase, and transferrin receptor IREs) and contrasts with the wild-type (WT) ferritin IRE containing an IL/B in protein binding assays (13). The ferritin Δ U6 IRE sequence thus is a useful model for comparisons of C-bulge and IL/B IREs, since other variations in sequence among natural IREs with a C-bulge will not be a factor. Included for comparison are data with full-length transcripts of ferritin mRNA containing the IRE hairpin loop (HL) mutations G18A or C14U/G18A (Figure 1A) that had previously been studied with a human growth hormone (hGH) reporter in transfected cells (23). Ferritin synthesis directed by the G18A-IRE ferritin full-length transcript, which had also previously been shown to be completely derepressed in rabbit reticulocyte lysates (RRL) (24), made a useful control for IRE-independent experimental variations.

Translation conditions were selected so that the rates of ferritin synthesis were linearly dependent on incubation time and mRNA concentration. RRL was selected because it had previously been shown to regulate ferritin mRNA with trans factors (14, 21, 25–27), now known to be IRP1 and IRP2 (13). In addition, ferritin synthesis in wheat germ extracts was measured to assess possible general effects of the mutations on protein synthesis, since wheat germ extracts have no ferritin mRNA-specific regulation (no IRPs) (21, 25, 28).

The results showed that all mutations in the IRE region studied increased ferritin synthesis in RRL, suggesting IRE-dependent derepression of translation (Figure 1B). However, deletion of U6 in the IL/B region or C14U/G18A double mutations in the HL increased ferritin synthesis by 4-fold

Table 1: IRE-Dependent Translation Repression and IRP Binding to the IRE

ferritin mRNA	repression in RRL ^a	binding to the IRE in ferritin mRNA	
		IRP1 ^b	IRP2 ^b
WT	94 \pm 3	100	100
Δ U6	66 \pm 3	87 \pm 5	<5
C14U/G18A	64 \pm 6	89 \pm 4	<5
G18A	0 ^c	<1	<1

^a Ferritin synthesis in RRL was analyzed by electrophoresis (Figure 1B) and quantitated by use of a PhosphorImager and ImageQuant software. The data are average of three experiments with two independent preparations of RNA in vitro transcripts. Repression = (1 - ferritin synthesis/G18A transcript-directed ferritin synthesis) \times 100.

^b Data from ref 13. ^c Repression for G18A ferritin mRNA was set to be zero, as neither IRP1 nor IRP2 binds the IRE in the mRNA, and was used to calculate repression for other tested ferritin mRNAs.

while G18A IRE mutation in the HL caused a 13-fold increase in translation. In wheat germ extract, which has no IRE recognition proteins (25, 28), all the transcripts tested were identical in directing ferritin synthesis (Figure 1C), showing that the differences in ferritin synthesis in RRL were dependent on IRP-regulated protein synthesis.

To facilitate quantitative comparisons of the effects of IRE structure on ferritin mRNA function, the data were normalized to those obtained with mRNA containing the G18A IRE, since neither IRP1 nor IRP2 is bound by the G18A IRE (13, 29–31) and ferritin synthesis is fully derepressed in RRL (24). The Δ U6 transcript was only repressed by 66% \pm 3%, whereas the repression for the WT transcript was 94% \pm 3% (Table 1). Repression of the C14U/G18A IRE transcript in RRL (64% \pm 6%) was comparable to the Δ U6 IRE transcript (Table 1) and confirms the results obtained with the same mutation and a hGH reporter transfected into cultured cells (23).

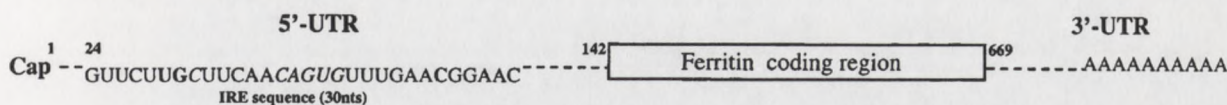
Effects of Internal Loop/Bulge and Hairpin Loop Mutations on Thermal Stability. The helix distortion, C-bulge, or internal loop/bulge (IL/B) in the iso-IREs influences both IRP2 recognition (13) and translation repression (Figure 1 and Table 1). The main structural difference between these two types of IREs is the number of nucleotides involved in the distortion at the midpoint of the helix (Figure 2). To explore the structural variation in the helix distortion and to examine the relationship of the helix distortion to IRE function, the thermal stability was analyzed for a series of IREs.

Conversion of the IL/B to a C-bulge by deletion of U6 increased the T_m and thermal stability (Table 2) as predicted from the change in secondary structures (Figure 2). Substitution of the hairpin loop (HL) CG base pair by UA (Figure 3) decreased T_m and thermal stability but had less effect than disruption of the base pair with a single mutation in the HL (G18A IRE) (Table 2), emphasizing the importance of the base pair in the HL to IRE stability. All these mutants, whether destabilizing or stabilizing IRE structures, decreased IRP2 binding (13) and translational regulation (Figure 1), suggesting that the RNA structure of the ferritin IRE is precisely balanced in an energy state associated with optimum protein recognition.

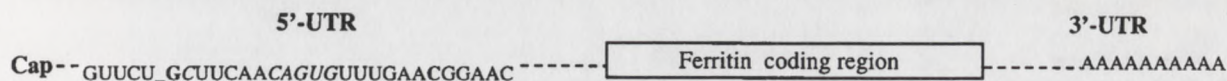
Effects of Internal Loop/Bulge and Hairpin Loop Mutations on Access to Nucleases: Cu(phen)₂ and RNase T1. Cu(phen)₂, a small reagent, was used to probe the structure of iso-IREs that displayed different binding of IRP2.

A

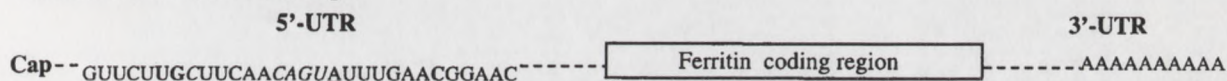
WT ferritin *in vitro* transcript:



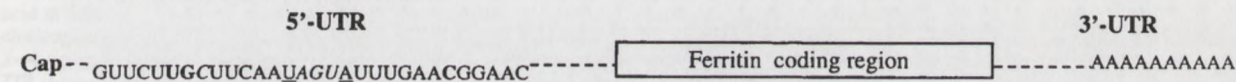
ΔU6 ferritin *in vitro* transcript:



G18A ferritin *in vitro* transcript:

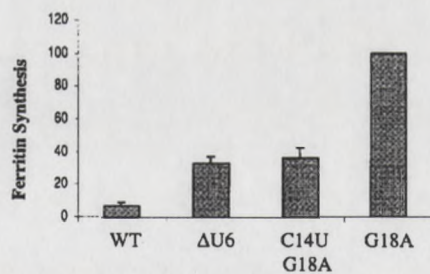
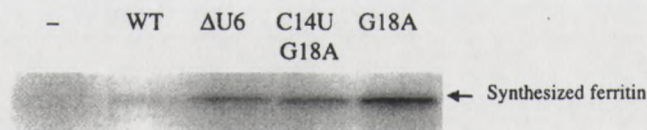


C14U/G18A ferritin *in vitro* transcript:



B

Transcript:



IRP1 binding :

IRP2 binding :

+	+	+	-
+	-	-	-

C

Transcript:

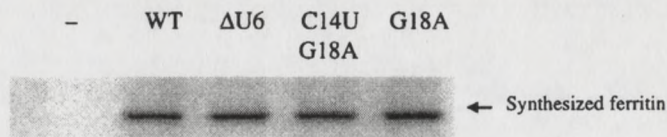


FIGURE 1: Effects of the IRE helix distortion (IL/B or C-bulge) on ferritin synthesis. Capped, polyadenylated full-length transcripts with the IL/B, C-bulge, and HL mutant IRE, schematically presented in panel A, were translated in rabbit reticulocyte lysates (B) or in wheat germ extracts (C) with [³⁵S]methionine. Ferritin synthesis was analyzed by electrophoresis in an SDS-polyacrylamide gel and with a PhosphorImager (Molecular Dynamics). The gel picture shown (B, C) is representative of three experiments with two independent preparations of the RNA transcripts. The wild-type and mutant ferritin IRE sequences are shown in panel A. *Italic* type indicates conserved residues in iso-IREs; **boldface** type indicates ferritin IRE-specific residues; underlined residues are mutant residues/sites in the 5' untranslated region.

W. T. ...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

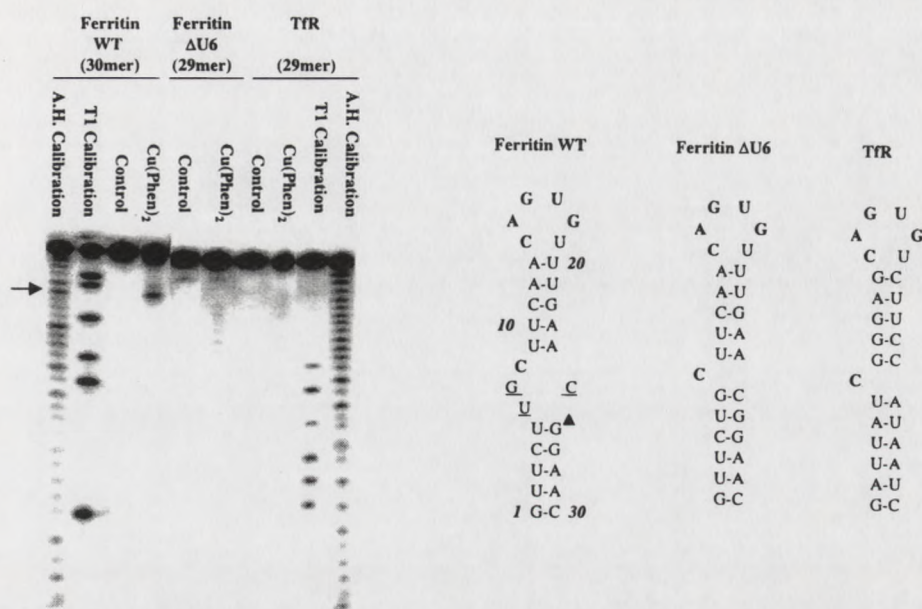


FIGURE 2: Effects of the IRE helix distortion (IL/B or C-bulge) on reactivity with $\text{Cu}(\text{phen})_2$ complexes. $5'$ - ^{32}P Labeled wild-type and mutant IREs (sequences and secondary structures on the right of the figure) were heated and slowly annealed as described under Materials and Methods. The annealed RNAs were incubated with or without $\text{Cu}(\text{phen})_2$, followed by initiation of the cleavage with 3-mercaptopropionic acid at 25°C and termination of the cleavage with 2,9-dimethyl-1,10-phenanthroline after 5 min. The cleavage products were resolved by electrophoresis and analyzed with a PhosphorImager and ImageQuant software (Molecular Dynamics). The experiment was repeated for 3–4 times with two independently prepared RNAs for each IRE. WT, wild type; ΔU6 , ferritin IRE with a single U deletion in the IL/B; TfR, transferrin receptor; A.H. calibration, alkaline hydrolysis of WT ferritin or TfR IRE; T1 calibration, RNase T1 cleavage of denatured WT or TfR IRE; control, without $\text{Cu}(\text{phen})_2$; $\text{Cu}(\text{phen})_2$, cleavage with $\text{Cu}(\text{phen})_2$. The arrow shows the site at which $\text{Cu}(\text{phen})_2$ cleavage was changed by engineered or natural mutations. The cleavage sites are also shown in the secondary structures (\blacktriangle); a significant cleavage site has $>10\times$ the intensity on average compared to insignificant cleavage sites. Note that the $\text{Cu}(\text{phen})_2$ reactivity was not significantly changed in the IRE with the C14U/G18A hairpin loop mutation (for clarity these data are not shown).

Table 2: Effects of Loop Mutations on Structures in Other Loops in the Ferritin IRE, Thermal Stability, and IRP2 Binding

IRE sequence ^a	IRP2 binding ^b	Reactivity ^c			T_m ^d ($^\circ\text{C}$)	ΔG_{37} ^d (kcal/mol)
		IL/B(G26/G27)		HL(G18)		
		$\text{Cu}(\text{phen})_2$	RNase T1	RNase T1		
WT	+	+	–	–	47.0 ± 0.7	-2.2 ± 0.2
HL C14U/G18A	–	+	+	na ^e	44.4 ± 1.0	-1.6 ± 0.1
HL G18A	–	+	+	na ^e	42.5 ± 0.9	-1.1 ± 0.1
IL/B ΔU6	–	–	–	–	52.3 ± 0.8	-3.8 ± 0.1
IL/B ΔU6 , ΔG7 ^f	nd ^g	–	–	+	nd ^g	nd ^g

^a HL; hairpin loop (CAGUGX; residues 14–18); IL/B; internal loop/bulge (UGC/C; residues 6, 7, 8/25). Note that the chemical shifts of imino proton resonances of loop residues U6, G7, G26 and G27, and G18 are all affected by Mg. RNase T1 reactivity toward the G residues in the HL is changed by IL/B mutation and reactivity toward the IL/B by HL mutation, emphasizing the structural interdependence of different parts of the ferritin IRE. Both types of changes eliminate IRP2 binding (13). ^b Data from (13). ^c See Materials and Methods and Figures 2 and 3. ^d Wild-type or mutant IRE ($\sim 1.5 \mu\text{M}$; for sequences and secondary structures, see Figures 2 and 3) in 1 mL of 10 mM phosphate buffer, pH 6.8, was thermally denatured between 5 and 95°C . The denaturation profile, absorbance at 260 nm as a function of temperature, was acquired on a Cary 100 spectrophotometer system. T_m and ΔG_{37} were determined from the profiles with hypochromicity methods (Cary Win Bio package application software). The data are the average of 3–4 experiments with two independently prepared RNA samples for each IRE; the error is presented as the standard deviation. Thermal denaturation of the RNAs in the phosphate buffer with 100 mM NaCl was also performed. For one experiment, the T_m (ΔG_{37}) for WT, ΔU6 , C14U/G18A, and G18A IREs are 55.0°C (-4.6 kcal/mol), 60.8°C (-6.1 kcal/mol), 52.5°C (-3.8 kcal/mol), and 50.5°C (-2.6 kcal/mol), respectively. ^e Not applicable, because the potential substrate, G, is replaced by A. ^f Data taken from ref 24. ^g Not determined.

Reactivity of a ferritin IRE with a single deletion in the internal loop/bulge (IL/B) (ferritin ΔU6 IRE) and the natural C-bulge IRE (transferrin receptor IRE) was compared to that of the wild-type (WT) ferritin IRE. The $\text{Cu}(\text{phen})_2$ complex is a small ($\sim 13 \text{ \AA}$ and 460 Da), structure-specific, sequence-independent probe that cleaves distortions in RNA helices (32–34) including the ferritin IRE (18, 19, 35).

The results showed that $\text{Cu}(\text{phen})_2$ cleaved the IRE 30-mer at G26 adjacent to the IL/B (Figure 2), as previously observed with natural mRNA or full-length ferritin in vitro

transcript (18, 19, 35). In contrast, no $\text{Cu}(\text{phen})_2$ cleavage was observed for either the engineered (ΔU6) or the natural (TfR) C-bulge IRE (Figure 2). Cleavage of the IRE was unaffected by the C14U/G18A Hairpin Loop (HL) mutation (data not shown in Figure 2 for clarity).

Thermal denaturation studies of WT and mutant ferritin IREs (Table 2) indicate that the GC base pair in the HL increases the overall stability by 50%. Given the diminished effect of mutating the G of the GC base pair in the HL on the stability of an IRE 16-mer (15%) (3), an effect of the

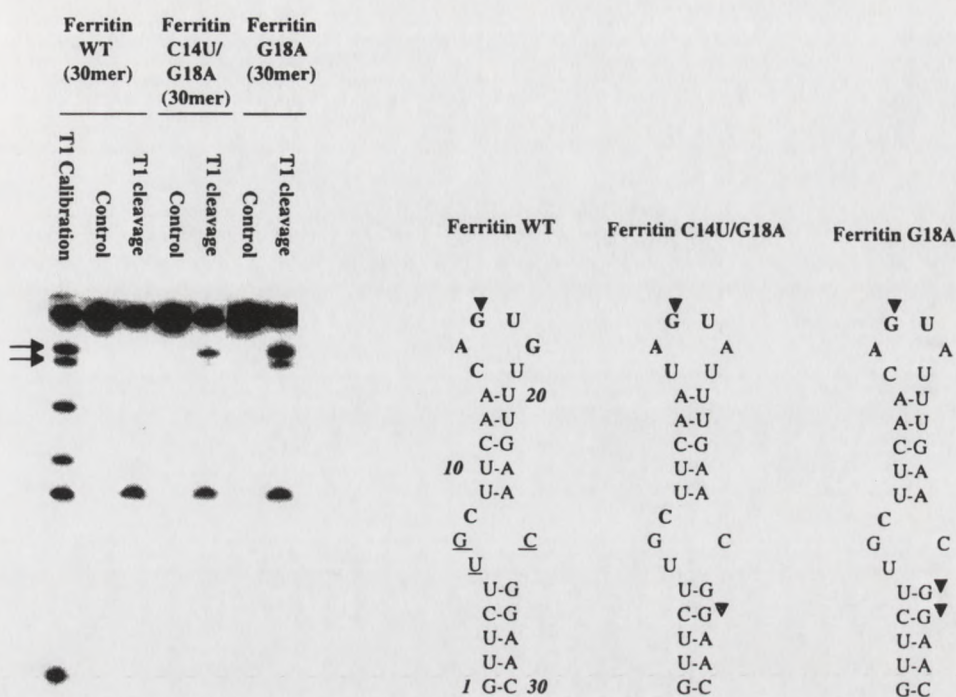


FIGURE 3: Effects of the hairpin loop mutations on RNase T1 cleavage at the IRE helix distortion. 5'-³²P-labeled wild-type and mutant IREs (sequences and secondary structures on the right of the figure) were heated and slowly annealed as described under Materials and Methods. RNase T1 cleavage of the annealed RNAs was accomplished by incubation of the RNAs with RNase T1 at 25 °C for 7 min. The cleavage products were resolved by electrophoresis and analyzed with a PhosphorImager and ImageQuant software (Molecular Dynamics). The experiment was repeated 3–4 times with two independently prepared RNAs for each IRE. WT, wild type; T1 calibration, RNase T1 cleavage of denatured WT IRE; control, incubation without RNase T1; T1 cleavage, incubation with RNase T1. Arrow shows the site in the IL/B at which RNase T1 cleavage was changed by mutation in the hairpin loop. The cleavage sites, shown in the secondary structures (▼), have >10× the intensity on average compared to insignificant cleavage sites. Note that the RNase T1 reactivity was unchanged in the ΔU6 IRE (for clarity these data are not shown).

HL mutation on the IL/B or helix structure is suggested. However, the small probe Cu(phen)₂ (~13 Å and 460 Da) did not detect structural changes in the HL mutant IRE.

RNase T1 is a larger probe (~30 Å and 11 000 Da) in contrast to the Cu(phen)₂ complex (32). Neither G26 nor G27 adjacent to the IL/B is cleaved by RNase T1 in the ferritin WT IRE (Figure 3) at room temperature (18, 19, 36), although G26 and G27 are cleaved by the smaller Cu(phen)₂ probe in the IRE 30-mer (Figure 2) and ferritin mRNA (18, 19). [Note that at high temperatures (45–50 °C), the G residues (G7 and G18) in the GC trans-loop base pairs are cleaved by RNase T1, revealing differential thermal stability compared to other regions of the helix (18, 19, 36).] The data in Figure 3 show that mutations in the HL at residues 14 and 18, which decreased IRP2 binding (5), translation repression (Figure 1), and *T_m* (Table 2), also changed structure around the IL/B, even though the sites of the mutations are more than 22 Å away from the IL/B (5). The mutations increased the sensitivity of G26 and G27 to RNase T1.

Effect of Mg²⁺ on the ¹H NMR spectrum of the Iron-Responsive Element. Previous NMR data and molecular modeling of Co(III) hexaammine binding to ferritin IRE indicated that the internal loop/bulge (IL/B) is a Co(III) hexaammine binding site (5). In addition, the IL/B is very sensitive to mutation and pH (Figures 2 and 3) (5, 24). Since Mg²⁺ is a physiological regulator of RNA structure and since both Co(III) hexaammine and Mg hexahydrate can bind at similar sites in other RNAs (37, 38), the possibility that Mg

binds in the IL/B was explored with one-dimensional NMR spectroscopy.

The ¹H NMR spectrum of the ferritin IRE with or without Mg²⁺ is shown in Figure 4A. Quantitation of Mg-induced change of chemical shifts of IRE imino proton resonances is shown in Figure 4B. Mg²⁺ produced a significant change of the chemical shifts of imino proton resonance for G26, G27, U5, U6, and G7 (Figure 4), which are all located in the IL/B, as did Co(III) hexaammine. However, Mg also induced significant chemical shifts for residues U10 and U21, in the helix between the IL/B and the hairpin loop (HL), and for G18, part of the base pair across the HL (Figure 4). The same sites for which Mg shifted the proton resonances were affected by Mg when Cu(phen)₂ cleavage was used to probe the structure of the IRE in natural poly(A)⁺ ferritin mRNA (18, 35). The results could reflect charge effects of Mg bound to the RNA complex or could indicate a structural change in the RNA.

DISCUSSION

A range of responses to cytoplasmic signals, such as iron, is possible among the mRNAs that contain IREs (11, 12, 39) despite the similarity of the iso-IRE secondary structure. All iso-IREs have the hairpin hexaloop (HL) with the conserved sequence CAGUGX and the helix distortion with a disordered C residue (4, 5, 7, 40–42). The range of physiological responses to iron among the mRNA with iso-IREs coincides with differences in binding of IRP2 (13). Two

The first part of the paper is devoted to a review of the literature on the topic. It is found that there is a general consensus that the model is well specified and that the data are well suited to the analysis. The second part of the paper is devoted to a description of the model and the data. The third part of the paper is devoted to a description of the estimation procedure. The fourth part of the paper is devoted to a description of the results. The fifth part of the paper is devoted to a description of the conclusions.

The first part of the paper is devoted to a review of the literature on the topic. It is found that there is a general consensus that the model is well specified and that the data are well suited to the analysis. The second part of the paper is devoted to a description of the model and the data. The third part of the paper is devoted to a description of the estimation procedure. The fourth part of the paper is devoted to a description of the results. The fifth part of the paper is devoted to a description of the conclusions.

The first part of the paper is devoted to a review of the literature on the topic. It is found that there is a general consensus that the model is well specified and that the data are well suited to the analysis. The second part of the paper is devoted to a description of the model and the data. The third part of the paper is devoted to a description of the estimation procedure. The fourth part of the paper is devoted to a description of the results. The fifth part of the paper is devoted to a description of the conclusions.

The first part of the paper is devoted to a review of the literature on the topic. It is found that there is a general consensus that the model is well specified and that the data are well suited to the analysis. The second part of the paper is devoted to a description of the model and the data. The third part of the paper is devoted to a description of the estimation procedure. The fourth part of the paper is devoted to a description of the results. The fifth part of the paper is devoted to a description of the conclusions.

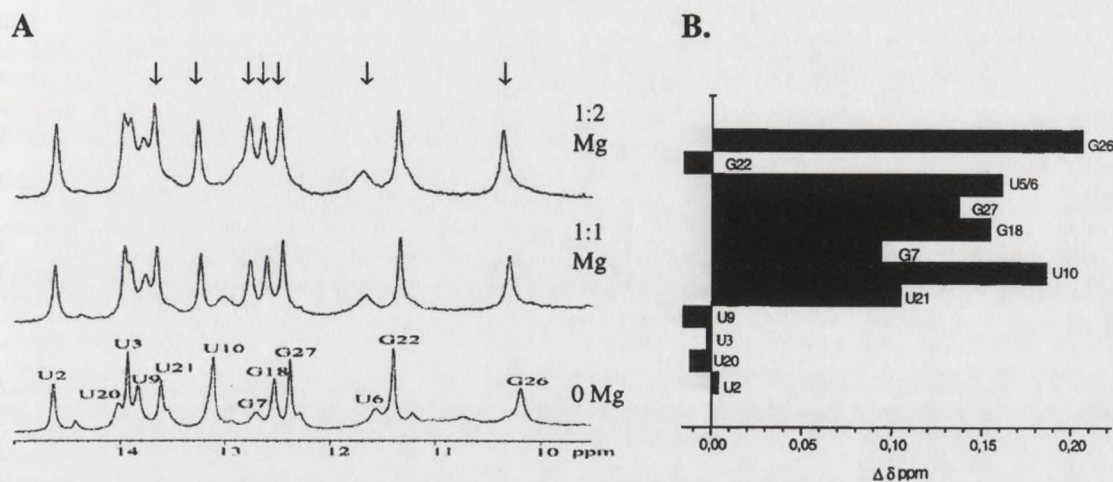


FIGURE 4: Effects of Mg^{2+} on chemical shifts of IRE imino resonances. The 1H NMR spectrum of the ferritin IRE (0.5 mM RNA) was collected at 12 °C in 10 mM sodium phosphate and 0.1 mM EDTA, pH 6.8, with or without Mg^{2+} . (A) 1H NMR spectra. Arrows show chemical shifts of imino proton resonances significantly upshifted by Mg (see panel B). The chemical shifts are for the residues in the IL/B, the HL, or the stem between them; no significant change of the chemical shifts for the residues in the lower stem was observed. (B) Quantitative comparison of Mg-induced shift of chemical shifts of imino proton resonances.

types of helix distortions known in the iso-IREs are a C-bulge (4) and an internal loop/bulge (IL/B) that contains the C residue. In the ferritin IRE, the IL/B forms a pocket that appears to bind Co(III) hexaammine and has a pH-dependent conformation (5). Coincidentally, the ferritin mRNA, which contains the IL/B iso-IRE, has the greatest range of iron regulation of any of the IRE-containing mRNAs both in vitro and in vivo (11, 12, 43–47). Even in an IL/B mRNA constructed from a ferritin IRE plus a human growth hormone (hGH) sequence, the ferritin IRE is more efficiently regulated than the transferrin receptor IRE (C-bulge) in the comparable hGH construct (48). In addition to the IL/B or C-bulge, the stem sequences among IREs in different mRNAs can diverge as much as 65% (7). Recently, differences in IRE/IRP interaction in vitro (13) were shown to coincide with the wide range of iron responses by IRE-containing mRNAs previously observed in vivo (11).

A connection between the different iron responses of IRE-containing mRNAs in rat liver (11) and the differential binding of IRP-1 and IRP-2 to short iso-IREs (29 and 30 nucleotides) in vitro (13) is shown by the correlation between translational repression in vitro of full-length, poly(A)+ mRNAs, IRE/IRP binding, the iron-induced rates of protein synthesis in vivo, and the presence of an IL/B or C-bulge in the IRE (Figure 1) (11–13, 43–48). IRP2 binding to the IRE is sensitive to the natural or engineered variation in the helix distortion (13) or to the engineered variation in the hairpin loop (29–31).

Chemical and enzymatic probing of IREs connects IRE structure with function measured as IRP2 binding and translation regulation (Table 2). Whether mutations were made in the HL or the IL/B, all mutations tested that displayed diminished IRP2 binding or translation repression also altered IL/B sensitivity to nuclease attack. For example, RNase T1 showed that the HL mutations made changes reflected in localized helix distortions near the IL/B 22 Å away (Figure 3 and Table 2), while the small probe $Cu(phen)_2$ detected localized effects of mutation in the IL/B itself (Figure 2 and Table 2). ^{31}P NMR spectra of wild-type

(WT) and HL mutant (G18A) IREs from previous NMR analysis (6) also indicated global perturbations in the helix backbone. Effects of the IL/B mutation on the HL were less dramatic than the effect of HL mutations on the IL/B and could not be detected with RNase T1 for the single deletion (Table 2). However, in a double IL/B deletion mutant ($\Delta U6 + \Delta G7$), both RNase T1 (Table 2) and $[Ru(tpy)(bpy)]$ (24) showed increased access to the HL. Clearly, the mutation in either loop region of the ferritin IRE influences the other loop and emphasizes interdependence of structural features throughout the ferritin IRE. [Preliminary data, (Ke and Theil, unpublished results) show that the thermal denaturation of the ferritin IRE is much more cooperative than that of any of the other IREs.]

Both IRE loops, the HL and IL/B, contribute to the specificity of IRP2 binding and translation repression (Figure 1 and Table 1), but the effects on the properties of the helix distortion around the conserved C residues differ. For example, disruption of the CG base pair or replacement of it with UA in the HL decreased the base stacking in the IRE (decreased T_m) (Table 2) and increased the helix distortion around the IL/B (increased sensitivity to RNase T1) (Figure 3). On the other hand, deletion of U6 in the IL/B increased base stacking in the IRE (increased T_m) (Table 2) and decreased the helix distortion at the IL/B [resistance to $Cu(phen)_2$ cleavage] (Figure 2). Thus, the CG trans-loop base pair in the HL appears to decrease distortions in the IRE, whereas the CG trans-loop base pair in the IL/B is associated with increased distortion in the helix near the conserved C residue. How the effects of Mg on IL/B, HL, and helix residues, detected by NMR spectroscopy (Figure 3) and $Cu(phen)_2$ cleavage (18), relate to IRE function has not yet been determined, but they may contribute to kinetics of protein binding. The negative impact of either destabilizing the IRE (HL mutations) or stabilizing the IRE (IL/B mutations) on IRP2 binding and translation repression (Figure 1 and Table 2) indicates that the native ferritin IRE structure (HL + IL/B) has the optimum stability and site accessibility for IRP2 binding and translation regulation. Iso-IRE RNA isoform

Handwritten notes or signatures, possibly including the name "M. J. ...".

Main body of the document containing several paragraphs of text, which is extremely faint and illegible.

variation, illustrated by the distinctive properties of the ferritin IRE, provides a natural set of mRNA-specific regulatory sequences. The combinatorial RNA/protein interactions possible with iso-IREs and iso-IRPs yield a range of mRNA responses to a single environmental signal.

REFERENCES

- Darimont, B. D., Wagner, R. L., Apreletti, J. W., Stallcup, M. R., Kushner, P. J., Baxter, J. D., Fletterick, R. J., and Yamamoto, A. R. (1998) *Genes Dev.* 12, 3343–3356.
- Alen, C., and Sonenshein, A. L. (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96, 10412–10417.
- Liang, L. G., and Hall, K. B. (1996) *Biochemistry* 35, 13586–13596.
- Address, K. J., Basilion, J. P., Klausner, R. D., Rouault, T. A., and Pardi, A. (1997) *J. Mol. Biol.* 274, 72–83.
- Gdaniec, Z., Sierzputowska-Gracz, H., and Theil, E. C. (1998) *Biochemistry* 37, 1505–1512.
- Sierzputowska-Gracz, H., McKenzie, R. A., and Theil, E. C. (1995) *Nucleic Acids Res.* 23, 145–152.
- Theil, E. C. (1998) in *Metal Ions in Biological Systems. Iron Transport and Storage in Microorganisms, Plants, and Animals* (Sigel, A., and Sigel, H., Eds.) pp 403–434, Marcel Dekker, Inc., New York.
- Eisenstein, R. L., Tuazon, P. T., Schalinske, K. L., Anderson, S. A., and Traugh, J. A. (1993) *J. Biol. Chem.* 268, 27363–27370.
- Schalinske, K. L., and Eisenstein, R. S. (1996) *J. Biol. Chem.* 271, 7168–7175.
- Beinert, H., Kennedy, M. C., and Stout, C. D. (1996) *Chem. Rev.* 96, 2335–2373.
- Chen, O. S., Schalinske, K. L., and Eisenstein, R. S. (1997) *J. Nutr.* 127, 238–248.
- Schalinske, K. L., Chen, O. S., and Eisenstein, R. S. (1998) *J. Biol. Chem.* 273, 3740–3746.
- Ke, Y., Wu, J., Leibold, E. A., Walden, W. E., and Theil, E. C. (1998) *J. Biol. Chem.* 273, 23637–23640.
- Dix, D. J., Lin, P.-N., Kimata, Y., and Theil, E. C. (1992) *Biochemistry* 31, 2818–2822.
- Dix, D. J., Lin, P.-N., McKenzie, A. R., Walden, W. E., and Theil, E. C. (1993) *J. Mol. Biol.* 231, 230–240.
- Fletcher, L., Corbin, S. D., Browning, K. S., and Ravel, J. M. (1990) *J. Biol. Chem.* 265, 19582–19587.
- Milligan, J. F., and Uhlenbeck, O. C. (1989) *Methods Enzymol.* 180, 51–62.
- Wang, Y.-H., Sczekan, S. R., and Theil, E. C. (1990) *Nucleic Acids Res.* 18, 4463–4468.
- Harrell, C. M., McKenzie, A. R., Patino, M. M., Walden, W. E., and Theil, E. C. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 1–6.
- Shull, G. E., and Theil, E. C. (1982) *J. Biol. Chem.* 257, 14187–14191.
- Dickey, L. F., Wang, Y.-H., Shull, G. E., Wortman, I. A., III, and Theil, E. C. (1988) *J. Biol. Chem.* 263, 3071–3074.
- Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* 2, 661–665.
- Menotti, E., Henderson, B. R., and Kuhn, L. C. (1998) *J. Biol. Chem.* 273, 1821–1824.
- Thorp, H. H., McKenzie, R. A., Lin, P.-N., Walden, W. E., and Theil, E. C. (1996) *Inorg. Chem.* 35, 2773–2779.
- Walden, W. E., Daniels-McQueen, S., Brown, P. H., Gaffield, L., Russell, D. A., Bielser, D., Bailey, L. C., and Thach, R. E. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 9503–9507.
- Leibold, E. A., and Munro, H. N. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 2171–2175.
- Bhasker, C. R., Burgiel, G., Neupert, B., Emery-Goodman, A., Kuhn, L. C., and May, B. K. (1993) *J. Biol. Chem.* 268, 12699–12705.
- Rothenberger, S., Mullner, E. W., and Kuhn, L. C. (1990) *Nucleic Acids Res.* 18, 1175–1179.
- Henderson, B. R., Menotti, E., Bonnard, C., and Kuhn, L. C. (1994) *J. Biol. Chem.* 269, 17481–17489.
- Henderson, B. R., Menotti, E., and Kuhn, L. C. (1996) *J. Biol. Chem.* 271, 4900–4908.
- Butt, J., Kim, H.-Y., Basilion, J. P., Cohen, S., Iwai, K., Philpott, C. C., Altschul, S., Klausner, R. D., and Rouault, T. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 4345–4349.
- Theil, E. C. (1994) *New J. Chem.* 18, 435–441.
- Murakawa, G. J., Chen, C. B., Kuwabara, M. D., Nierlich, D. P., and Sigman, D. S. (1989) *Nucleic Acid Res.* 17, 5362–5375.
- Sigman, D. S. (1990) *Biochemistry* 29, 9097–9105.
- Wang, Y.-H., Sczekan, S. R., McKenzie, R. A., and Theil, E. C. (1991) *Biol. Met.* 4, 56–61.
- Bettany, A. J. E., Eisenstein, R. S., and Munro, H. M. (1992) *J. Biol. Chem.* 267, 16531–16537.
- Cate, J. H., and Doudna, J. A. (1996) *Structure* 4, 1221–1229.
- Kieft, J. J., and Tinoco, I. (1997) *Structure* 5, 713–721.
- Eisenstein, R. S., and Blemings, K. P. (1998) *J. Nutr.* 128, 2295–8.
- Theil, E. C. (1994) *Biochem. J.* 304, 1–11.
- Rouault, T. A., and Klausner, R. D. (1996) *J. Biol. Inorg. Chem.* 1, 494–499.
- Hentze, M. W., and Kuhn, L. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 8175–8182.
- Schaefer, F. V., and Theil, E. C. (1981) *J. Biol. Chem.* 256, 1711–1715.
- Shull, G. E., and Theil, E. C. (1983) *J. Biol. Chem.* 258, 7921–7923.
- Dandekar, T., Stripecke, R., Gray, N. K., Goossen, B., Constable, A., Johansson, H. E., and Hentze, M. W. (1991) *EMBO J.* 10, 1903–1909.
- Melefsors, O., Goossen, B., Johansson, H. E., Stripecke, R., Gray, N. K., and Hentze, M. W. (1993) *J. Biol. Chem.* 268, 5974–5978.
- Kim, H.-Y., LaVaute, T., Iwai, K., Klausner, R. D., and Rouault, T. A. (1996) *J. Biol. Chem.* 271, 24226–24230.
- Casey, J. L., Hentze, M. W., Koeller, D. M., Caughman, S. W., Rouault, T. A., Klausner, R. D., and Harford, J. B. (1988) *Science* 240, 924–928.

BI9924765

of the ...

REFERENCES

1. ...
2. ...
3. ...
4. ...
5. ...
6. ...
7. ...
8. ...
9. ...
10. ...
11. ...
12. ...
13. ...
14. ...
15. ...
16. ...
17. ...
18. ...
19. ...
20. ...
21. ...
22. ...
23. ...
24. ...
25. ...
26. ...
27. ...
28. ...
29. ...
30. ...
31. ...
32. ...
33. ...
34. ...
35. ...
36. ...
37. ...
38. ...
39. ...
40. ...
41. ...
42. ...
43. ...
44. ...
45. ...
46. ...
47. ...
48. ...
49. ...
50. ...
51. ...
52. ...
53. ...
54. ...
55. ...
56. ...
57. ...
58. ...
59. ...
60. ...
61. ...
62. ...
63. ...
64. ...
65. ...
66. ...
67. ...
68. ...
69. ...
70. ...
71. ...
72. ...
73. ...
74. ...
75. ...
76. ...
77. ...
78. ...
79. ...
80. ...
81. ...
82. ...
83. ...
84. ...
85. ...
86. ...
87. ...
88. ...
89. ...
90. ...
91. ...
92. ...
93. ...
94. ...
95. ...
96. ...
97. ...
98. ...
99. ...
100. ...

Part
of

Department
Chairman

Adm.
Asst.
Gen.
& LP
R. R.
P. P.
O. D.
Gen.
W. C.
Gen.
A. C.
M. C.
W. C.
Gen.
Gen.
Gen.
Gen.
Gen.
Gen.
Gen.
Gen.

V.

Lyons, W. C.
role in research
DNA
protection
mutational
antigen
terminus
adhesion
lymphocyte
application

The present
National
Service
Education
Research
To
The
The
The
The

.v

Functional Anticodon Architecture of Human tRNA^{Lys3} Includes Disruption of Intraloop Hydrogen Bonding by the Naturally Occurring Amino Acid Modification, t⁶A[†]

John W. Stuart,[‡] Zofia Gdaniec,[§] Richard Guenther,[‡] Michal Marszalek,^{||} Elzbieta Sochacka,^{||} Andrzej Malkiewicz,^{||} and Paul F. Agris^{*;‡}

Department of Biochemistry, North Carolina State University, Raleigh, North Carolina 27695-7622, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan 61-704, Poland, and Institute of Organic Chemistry, Technical University, Lodz 90-924, Poland

Received June 6, 2000; Revised Manuscript Received September 1, 2000

ABSTRACT: The structure of the human tRNA^{Lys3} anticodon stem and loop domain (ASL^{Lys3}) provides evidence of the physicochemical contributions of N⁶-threonylcarbamoyladenine (t⁶A₃₇) to tRNA^{Lys3} functions. The t⁶A₃₇-modified anticodon stem and loop domain of tRNA^{Lys3}_{UUU} (ASL^{Lys3}_{UUU}-t⁶A₃₇) with a UUU anticodon is bound by the appropriately programmed ribosomes, but the unmodified ASL^{Lys3}_{UUU} is not [Yarian, C., Marszalek, M., Sochacka, E., Malkiewicz, A., Guenther, R., Miskiewicz, A., and Agris, P. F., *Biochemistry* 39, 13390–13395]. The structure, determined to an average rmsd of 1.57 ± 0.33 Å (relative to the mean structure) by NMR spectroscopy and restrained molecular dynamics, is the first reported of an RNA in which a naturally occurring hypermodified nucleoside was introduced by automated chemical synthesis. The ASL^{Lys3}_{UUU}-t⁶A₃₇ loop is significantly different than that of the unmodified ASL^{Lys3}_{UUU}, although the five canonical base pairs of both ASL^{Lys3}_{UUU} stems are in the standard A-form of helical RNA. t⁶A₃₇, 3'-adjacent to the anticodon, adopts the form of a tricyclic nucleoside with an intrasidic H-bond and enhances base stacking on the 3'-side of the anticodon loop. Critically important to ribosome binding, incorporation of the modification negates formation of an intraloop U₃₃·A₃₇ base pair that is observed in the unmodified ASL^{Lys3}_{UUU}. The anticodon wobble position U₃₄ nucleobase in ASL^{Lys3}_{UUU}-t⁶A₃₇ is significantly displaced from its position in the unmodified ASL and directed away from the codon-binding face of the loop resulting in only two anticodon bases for codon binding. This conformation is one explanation for ASL^{Lys3}_{UUU} tendency to prematurely terminate translation and -1 frame shift. At the pH 5.6 conditions of our structure determination, A₃₈ is protonated and positively charged in ASL^{Lys3}_{UUU}-t⁶A₃₇ and the unmodified ASL^{Lys3}_{UUU}. The ionized carboxylic acid moiety of t⁶A₃₇ possibly neutralizes the positive charge of A⁺₃₈. The protonated A⁺₃₈ can base pair with C₃₂, but t⁶A₃₇ may weaken the interaction through steric interference. From these results, we conclude that ribosome binding cannot simply be an induced fit of the anticodon stem and loop, otherwise the unmodified ASL^{Lys3}_{UUU} would bind as well as ASL^{Lys3}_{UUU}-t⁶A₃₇. t⁶A₃₇ and other position 37 modifications produce the open, structured loop required for ribosomal binding.

Lysine tRNAs with UUU anticodons have a conventional role in ribosome-mediated protein synthesis. In addition, tRNA^{Lys}_{UUU} species facilitate -1 frameshifts for correct translation of the *E. coli* DNA polymerase γ subunit (1) and retroviral polymerases (2). Also, tRNA^{Lys}_{UUU} often misreads asparagine codons (3, 4) and peptidyl-tRNA^{Lys} prematurely terminates translation more often than other tRNAs (5). In addition, reverse transcription of the HIV-1 genomic RNA is primed by the human tRNA^{Lys}_{UUU}. Formation of the viral replication initiation complex is enhanced in vitro by the

presence of the tRNA's modified nucleosides (6, 7) and strand transfer is facilitated by the anticodon's modified nucleosides (8). We hypothesized that anticodon domain modified nucleosides impart the unique chemical and structural properties required to explain the standard (9), as well as the unconventional, roles of tRNA^{Lys}_{UUU} in protein synthesis and as primer for HIV replication (10).

Two posttranscriptional modifications in the anticodon loop distinguish tRNA^{Lys}_{UUU} from other tRNA^{Lys} species and from other tRNAs, in general. In human tRNA^{Lys}_{UUU}, position 34 is modified to 5-methoxycarbonylmethyl-2-thiouridine (mcm⁵s²U₃₄)¹ and A₃₇ is modified to 2-methylthio-N⁶-threonylcarbamoyladenine (ms²t⁶A₃₇). There is only one lysine tRNA in *E. coli*, tRNA^{Lys}_{UUU}, and its anticodon stem and loop sequence closely resembles that of human tRNA^{Lys}_{UUU}. In *E. coli* tRNA^{Lys}_{UUU}, U₃₄ and A₃₇ are modified to 5-methylaminomethyl-2-thiouridine (mnm⁵s²U) and t⁶A, respectively. Posttranscriptional modifications of

[†] This research was supported by the National Science Foundation (MCB9631103 to P.F.A.), the Department of Health and Human Services (PHS NIH Grant GM-23027 to P.F.A.), Cambridge Isotope Laboratories, Inc. (P.F.A.), and the Polish Committee for Scientific Research Grant 0877/T09/96/11 (A.M.).

* To whom correspondence should be addressed. Phone: (919) 515-6188. Fax: (919) 515-2047. E-mail: Agris@bchserver.bch.ncsu.edu.

[‡] North Carolina State University.

[§] Polish Academy of Sciences.

^{||} Technical University.

Functional Analysis of the ...

by ...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

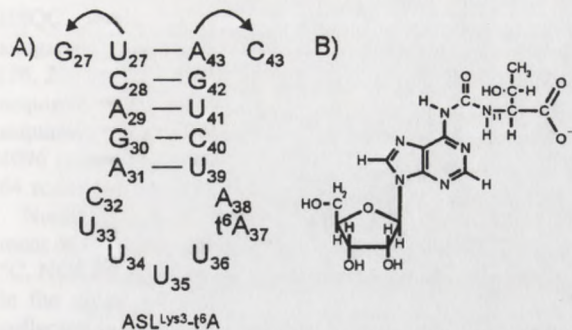


FIGURE 1: (A) Sequence and secondary structure of the tRNA^{Lys3} anticodon stem/loop (ASL^{Lys3}) with the t⁶A₃₇ modification. The molecule used for structure determination had the terminal base pair U₂₇·A₄₃ replaced with G₂₇·C₄₃. (B) The structure of t⁶A₃₇ with the carboxylic acid of threonine dissociated (pH 5.6).

the tRNA^{Lys} anticodon loop, but not those of the anticodon stem, are important for aminoacyl-tRNA synthetase recognition and aminoacylation of cognate tRNA (11) and for ribosomal binding (12, 13). Recently, we determined that the individual modifications of the tRNA^{Lys} anticodon loop, s²U₃₄, mnm⁵U₃₄ (of mnm⁵s²U₃₄) and t⁶A₃₇, restored AAA-programmed ribosomal binding to the otherwise unmodified and nonfunctional human tRNA^{Lys} anticodon stem and loop domain (13, 14, preceding paper in this issue).

Here we report the structure of the human tRNA^{Lys3} anticodon stem and loop with the amino acid modification, t⁶A₃₇ (ASL^{Lys3}-t⁶A₃₇), which is critical to its ribosome binding function (14). Structures of the unmodified ASL^{Lys3} and the stem modified ASL^{Lys3}-Ψ₃₉ have been reported (15), but neither bind the ribosome (14). We compare the structure of the functional ASL^{Lys3}-t⁶A₃₇ to that of the nonfunctional, unmodified ASL^{Lys3}. The hypermodified nucleoside has a pronounced effect on the anticodon loop conformation as determined from NMR-derived distance and torsion angle restraints using restrained molecular dynamics. The structure explains the role of t⁶A in achieving an anticodon architecture required for ribosome binding, as well as the unique properties displayed by tRNA^{Lys} species.

EXPERIMENTAL PROCEDURES

Materials: Sample Preparation. The heptadecamer oligoribonucleotides corresponding to the unmodified and N⁶-threonylcarbamoyladenine (t⁶A₃₇)-modified sequence of the human tRNA^{Lys3} anticodon stem and loop domain, ASL^{Lys3} (Figure 1) were chemically synthesized using standard phosphoramidite chemistry on an Applied Biosystems 394 DNA/RNA Synthesizer (16). The 5' to 3' terminal base pair was changed from the native sequence Ψ₂₇·A₄₃ to G₂₇·C₄₃ for the increased yield and increased stability necessary for NMR analysis. The 5'-trityl-2'-dimethyltert-

butylsilyl-3'-¹⁵N₃-uridine phosphoramidite was synthesized as previously described (17) and the protected 5'-trityl-2'-dimethyltertbutylsilyl-3'-phosphoramidite of t⁶A was prepared in our laboratories (unpublished). All other nucleoside phosphoramidites were obtained from Glen Research (Sterling, VA). The oligomer was HPLC-purified as previously described (17) using a Nucleogen 60-7 DEAE (250 × 10 mm) column, and its nucleoside composition determined (18). Sample desalting was accomplished with Waters Corporation Sep-pak columns. The sample was prepared for NMR spectroscopy by extensive dialysis with buffer (10 mM cacodylate buffer, pH 5.6, 0.1 mM EDTA, 6% D₂O/94% H₂O) using Amicon Centricon 3 concentrators. For experiments in D₂O, the sample solution was evaporated down with N₂ and exchanged with D₂O three times before resuspending the sample in 99.96% D₂O. The samples used for structural analysis had a final RNA concentration of 1.2 mM.

Methods: (i) ASL Thermodynamic Parameters. ASL^{Lys3} samples were dissolved to a concentration of 2 μM in either a phosphate or cacodylate buffer (10 mM sodium phosphate or 10 mM sodium cacodylate, 100 mM NaCl, 0.1 mM EDTA, with the pH adjusted with HCl). Thermal denaturations, performed in triplicate, were monitored by UV absorbance (260 nm) using a Cary 3 spectrophotometer as previously described (19). Data points were averaged over 20 s and collected three times a minute. Denaturations and renaturations were conducted over a temperature range of 5–90 °C and at three pH values (7.2, 6.0, and 5.0) with a ramp rate of 1 °C/min. Data from denaturations and renaturations were treated similarly. No hysteresis was observed. Thermodynamic parameters were calculated with a van't Hoff analysis of the data as described by Serra and Turner (20) using Origin software (Microcal). Substitution of the U₂₇·A₄₃ terminal base pair with a G₂₇·C₄₃ added stability to the molecule. A pH between 5.0 and 6.0 was deemed to be optimal for stability and pH 5.6 was chosen for NMR studies.

(ii) Unimolecular Property of Samples. To demonstrate that the RNA was a monomer at the NMR concentrations of this study, we measured the translational diffusion coefficient of the unmodified and t⁶A₃₇-modified ASLs, and compared them to that of other RNA sequences of various lengths. This was accomplished with the pulsed field-gradient spin-echo technique (21, 22), performed on the NMR sample itself, thereby obviating any complications or ambiguities of interpretation that may arise from approaches other than that by NMR methods (23). As would be expected for a monomer of this size, the translational diffusion coefficient of the heptadecamer was bracketed by those measured for a hexamer and a dodecamer, and by that of a 28-mer. Our results are in good agreement with published values of the translational diffusion coefficients of nucleic acids of similar sizes (23, 24).

(iii) NMR Spectroscopy. All NMR spectra were collected on a Bruker DRX500 spectrometer and processed using either XWINNMR (Bruker Inc., Rheinstetten, Germany) or FELIX (Biosym/MSI, San Diego, CA). The residual water peak in D₂O samples was suppressed using low power presaturation whereas exchangeable proton resonances for samples in 94% H₂O/6% D₂O were collected at 1 °C with WATERGATE (25) solvent suppression. Eight one-dimensional spectra as a function of temperature between 4 and 31 °C and three

¹ Abbreviations: ASL, anticodon stem and loop domain; t⁶A, N⁶-threonylcarbamoyladenine; mcm⁵s²U, 5-methoxycarbonylmethyl-2-thiouridine; Ψ, pseudouridine; NMR, nuclear magnetic resonance; HPLC, high-performance liquid chromatography; DEAE, diethylaminoethyl; EDTA, ethylenediaminetetraacetic acid; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; HETCOR, heteronuclear correlation; HSQC, heteronuclear single quantum coherence; DQ, double quantum; DQF-COSY, double quantum filtered correlation spectroscopy; rmsd, root-mean-square deviation.



Figure 1

The diagram illustrates the structure of a multi-layered system, showing the arrangement of layers and the direction of flow or interaction. The layers are numbered 1 through 6, with layer 1 at the top and layer 6 at the bottom. The flow is indicated by arrows pointing downwards from the top layer.

The text describes the properties and interactions of these layers, including the role of various components and the overall system's behavior. It discusses the relationship between the layers and the external environment, as well as the internal dynamics of the system.

The system is characterized by its layered structure, which allows for the separation and interaction of different components. The flow of material or information is controlled by the arrangement of these layers, and the system's response is determined by the interactions between them.

The diagram and text provide a detailed view of the system's architecture and its operational principles. The layers are designed to facilitate specific processes, and the overall structure is optimized for efficiency and control.

The following text provides further details on the system's components and their interactions. It discusses the specific roles of each layer and how they contribute to the overall function of the system.

The system's performance is influenced by the thickness and composition of the layers, as well as the nature of the interactions between them. The text explores these factors and their impact on the system's behavior.

The diagram and text together provide a comprehensive understanding of the system's structure and function. The layered design is a key feature that enables the system to perform its intended tasks effectively.

The system's design is based on the principles of layering and interaction, which are fundamental to its operation. The text and diagram illustrate how these principles are applied in the system's architecture.

HSQC spectra at three pH values were obtained to aid signal assignments and assess protonation of A₃₈. NOESY spectra (26, 27) used for analysis of the exchangeable protons were acquired with a mixing time of 150 ms. Spectra were acquired with sweep widths of 12019 Hz in both dimensions, 4096 points in *t*₂, and a minimum of 512 points in *t*₁ with 64 scans per increment.

Nonexchangeable protons were assigned using an assortment of two-dimensional spectra acquired at 10, 16, and 25 °C. NOESY spectra of the sample in D₂O had mixing times in the range of 50 to 400 ms. At least 256 points were collected in *t*₁, with 64 scans per increment. For these and the following homonuclear experiments, the spectral width was 6000 Hz in both dimensions, and 1024 points were collected in *t*₂. TOCSY (28) or clean TOCSY (29) experiments using the MLEV17 mixing sequence were performed at 16 °C with a mixing time of 60 ms. DQF-COSY (30) and double quantum (DQ) (31, 32) experiments were obtained at 16 °C. The two-dimensional heteronuclear spectra that were collected included a ¹H-³¹P HETCOR (33) and a hetero-TOCSY-NOESY (34). These spectra were all acquired at 16 °C. Natural abundance ¹H-¹³C HSQC experiments (35-37) were performed at both 16 and 25 °C. For samples with site-specific ¹⁵N labels, ¹H-¹⁵N HSQC spectra were obtained in H₂O at 1 °C.

Distance restraints were derived from a series of phase sensitive NOESY spectra acquired with a 5000 Hz spectral width in both dimensions, 2048 points in *t*₂ and 512 points with 64 scans per block in *t*₁. Spectra were collected at 16 °C with mixing times of 50, 100, 140, 200, and 400 ms and processed (XWINNMR) with 60° phase-shifted sine bell apodization functions. The baseline in both dimensions was treated in FELIX using the FLATT algorithm (38). To provide suitable digital resolution for cross-peak integration, the spectra were zero-filled to 2048 by 1024 points.

(iv) *Structure Determination.* Distance restraints between nonexchangeable protons were obtained from the NOESY mixing time study. The NOE buildup curves were calculated by integrating the cross-peaks using FELIX and normalized by setting pyrimidine H5-H6 cross-peaks to a distance of 2.44 Å. Upper and lower bounds were set to the 20% above or below the calculated distance using flat-bottomed quadratic potentials. Cross-peaks with more than 35% overlap were classified strong (1.8-4 Å), medium (1.8-5 Å), or weak (1.8-6 Å) with large bounds to account for ambiguity of the peaks volume measurement. Threonine methyl protons of t⁶A were handled as pseudo atoms and adjusted accordingly (39). The distance restraints involving exchangeable resonances were obtained from a single NOESY spectrum of the sample in 94% H₂O/6% D₂O. The cross-peaks were qualitatively classified as strong (1.8-3.5 Å), medium (1.8-4.5 Å), or weak (1.8-5.5 Å).

Dihedral angle restraints were placed on the δ backbone (C5'-C4'-C3'-O3') to characterize the sugar puckers based on the ³J_{H1'H2'} from the DQF-COSY spectrum. Observable H1'-H2' cross-peaks with coupling constants less than 3 Hz were constrained to the C3'-endo conformation (δ = 85 ± 30°) while those with coupling constants greater than 7 Hz were constrained to the C2'-endo conformation (δ = 160 ± 30°). Sugars with intermediate coupling constants were left unrestrained. The α and ζ torsion angles were loosely constrained to exclude the trans conformation for those

Table 1: Thermodynamic Parameters^a for ASL^{Lys3} Constructs

ASL construct	pH	T _m (°C)	ΔG ₃₇ (kcal/mol) ^b	ΔH (kcal/mol)	ΔS (cal/mol K)
ASL ^{Lys3} _{UUU} unmodified	7.2 ^c	57.5	-3.4	-55.4	-168
	6.0	61.6	-4.2	-57.8	-172
	5.0	56.3	-3.5	-60.6	-184
ASL ^{Lys3} _{UUU-t⁶A₃₇}	7.2 ^c	54.7	-2.7	-49.5	-151
	6.0	60.2	-3.5	-50.6	-151
	5.0	58.4	-3.1	-48.0	-143

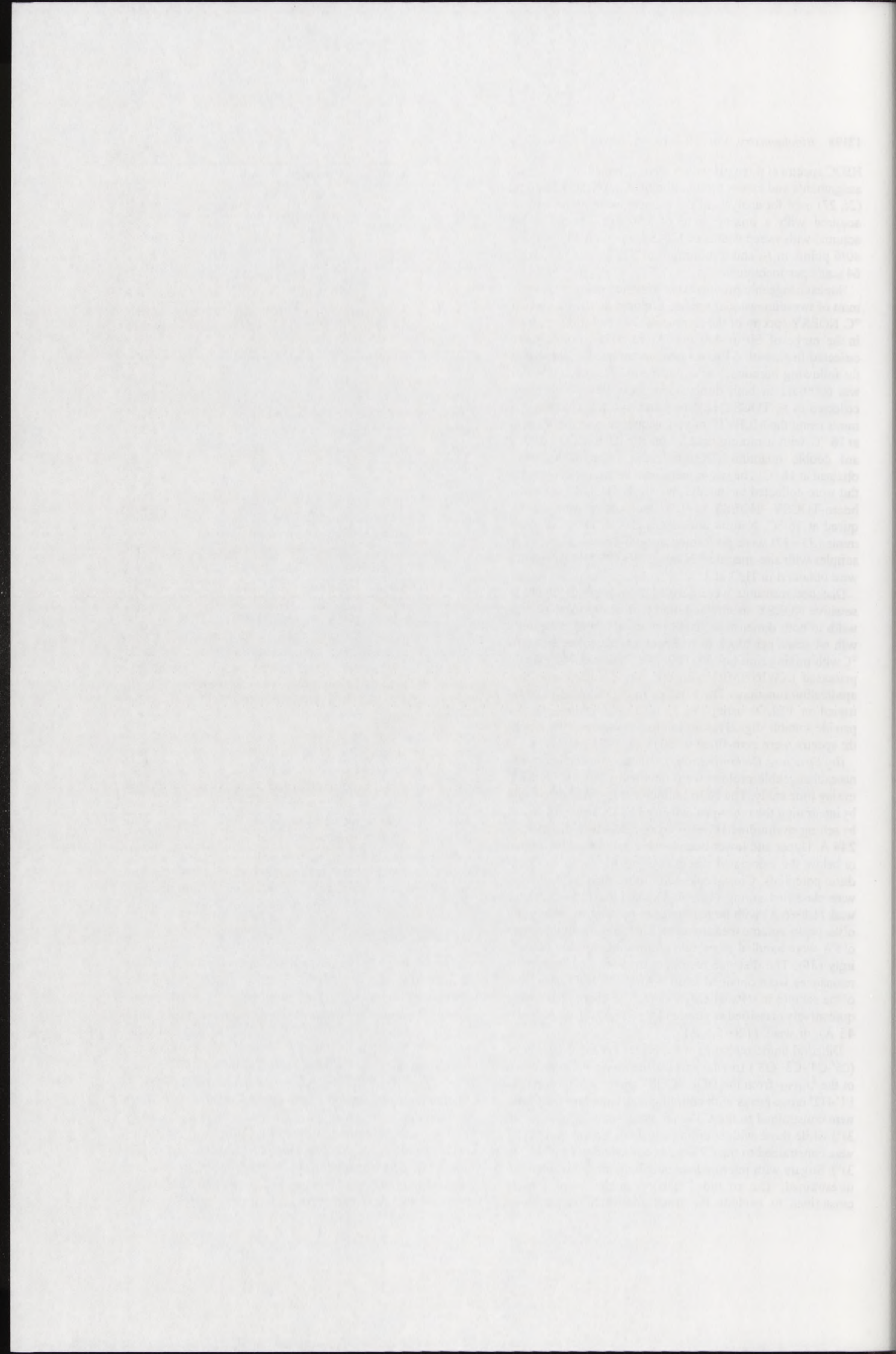
^a Error in determinations: T_m, ±0.9 °C; ΔG₃₇, ±0.4; ΔH, ±3.0; ΔS, ±7. ^b ΔG₃₇ determined at 37 °C. ^c Also cited in the preceding paper in this issue (14).

residues whose ³¹P chemical shifts fell within the narrow range commonly seen for regular A-form structures (40, 41). Last to better define the hydrogen-bonding pattern, distance restraints were added to the five base pairs in the stem and through-space dihedral restraints (±10°) were added to maintain planar bases.

Structure calculations were performed using InsightII (MSI) and protocols of Varani and others (40, 42). To achieve the global fold of the molecule, 50 initial distance geometry structures were generated from a matrix of random trial distances derived from covalent bonds, distance, dihedral angle, and chiral restraints. The resulting structures were regularized as part of the distance geometry protocol by simulated annealing using the default settings in InsightII. For the refinement phase, the AMBER force field (43) was used with the standard nucleoside parameters except for the modified nucleoside t⁶A₃₇ and adenosine protonated at the 1 position (A⁺₃₈). The state of t⁶A at pH 5.6 with the carboxylic acid dissociated (44-46) was used to determine its partial charges, which were calculated from scaled MOPAC charges (47, 48). Its atom types were derived by merging the existing AMBER atom types for adenine with that of threonine (43). The set of 50 distance geometry structures was heated to 2000 K for 5 ps with a 0.5 fs time step. During this equilibration step, the dihedral and chiral restraints (*k*_{dihedral} = 50 kcal mol⁻¹ rad⁻², *k*_{chiral} = 25 kcal mol⁻¹ rad⁻²) were scaled from 10% to their full value while the distance restraints remained at 50 kcal mol⁻¹ Å⁻². Next, the structures were cooled from 2000 to 100 K over 25 ps with a time step of 0.5 fs during which all restraints were maintained at their full values. Longer simulation times did not improve the convergence of the lowest energy structures.

RESULTS

Thermodynamic Parameters. The anticodon stem and loop domain of human tRNA^{Lys3}_{UUU}, ASL^{Lys3}_{UUU} (Figure 1), was chemically synthesized with and without the threonine-modified adenosine N6-threonylcarbamoyl-adenosine, t⁶A₃₇. At pH 7.2, the unmodified ASL^{Lys3}_{UUU} had a significantly higher melting temperature (T_m = 57.5 °C) and enhanced overall stability (ΔG = -3.4 kcal/mol) than the ASL^{Lys3}_{UUU-t⁶A₃₇} (T_m = 54.7 °C and ΔG = -2.7 kcal/mol) (Table 1). Interestingly, the introduction of t⁶A₃₇ made a significant contribution to ΔS, though the T_m of the RNA had decreased. The two RNAs exhibited greater stability at pH 6 than at pH 7.2 or pH 5.0 (Table 1 and ref 14). The increased stability and NMR-derived solution structures of the unmodified and Ψ₃₉-containing heptadecamers at low pH had already been reported and attributed to a C₃₂·A⁺₃₈ base pair (15). At the



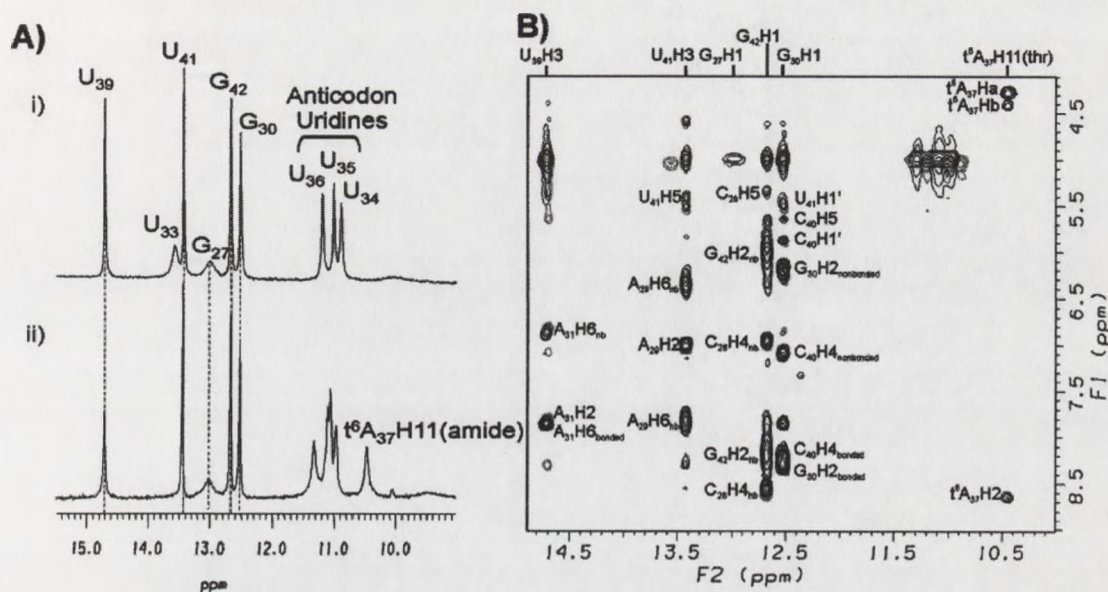


FIGURE 2: Comparison of exchangeable proton spectra of ASL^{Lys3}_{UUU}-t⁶A₃₇ with that of the unmodified ASL^{Lys3}_{UUU}. (A) One-dimensional ¹H spectra in H₂O of ASL^{Lys3}_{UUU}-t⁶A₃₇ (i) and unmodified ASL^{Lys3}_{UUU} (ii). Iminos involved in base pairing in the stem are shown with dashed lines. (B) Overlay of the two-dimensional H₂O NOESY spectra of the imino region with ASL^{Lys3}_{UUU}-t⁶A₃₇ shown in black and the unmodified ASL^{Lys3}_{UUU} shown in red.

lower pH 5.0, protonation of Cs destabilizes the RNAs. Thus, we determined the ASL^{Lys3}_{UUU}-t⁶A₃₇ structure at pH 5.6 by NMR spectroscopy and restrained molecular dynamics.

Assignment of Exchangeable Resonances. The one-dimensional and NOESY spectra of the unmodified and modified ASLs in H₂O revealed that the ASL^{Lys3}_{UUU}-t⁶A₃₇ was stable and had a single predominant conformation (Figure 2). Sequential imino-imino connectivities for the stem were observed, indicating that the RNA adopted a folded structure. As is often the case, however, due to fraying, these sequential connectivities did not extend to the terminal base pair, G₂₇·C₄₃. The broad signal of the base-paired G₂₇ was visible in both the one-dimensional and NOESY spectra (Figure 2). In addition to the imino resonances of the stem, t⁶A₃₇ H11, corresponding to the amide NH of threonine, had cross-peaks to the amino acid's H α and H β and the purine's H2 (Figure 2B). This is the pattern one would expect with a side-chain coplanar to the parent base (49, 50) and results in the modified nucleoside adopting the form of a tricyclic base. When the NOESY spectrum of the imino to amino region of the ASL^{Lys3}_{UUU}-t⁶A₃₇ was compared to that of the unmodified sequence, a very similar pattern of resonance cross-peaks (except that contributed by t⁶A) was observed with very little change in the chemical shifts. Likewise, comparison of other parts of the spectrum showed small changes in the chemical shifts of the resonances assigned to the stem, but considerably larger differences to the loop resonances. Cross-peaks observed in the imino to imino, and imino to amino, aromatic, and H1' regions of the spectra were those expected of right-handed, helical nucleic acids (35, 51). This implies that the t⁶A₃₇ modification did little to perturb the stem of the hairpin, but significantly affected the conformation of the loop region. At pH 5.6, the N1 of A₃₈ is protonated in the unmodified ASL^{Lys3}_{UUU} (15). Though this imino proton may be involved in a hydrogen bond with C₃₂ and thereby increase stability as was observed in thermal

denaturation studies at pH 6 compared to 7.2, it was not directly observable by NMR. However, a pH dependent change in the chemical shift of the A₃₈H2, indicative of the protonation of the base (15), was observed in HSQC (¹H-¹³C) spectra of both the unmodified ASL^{Lys3}_{UUU} and ASL^{Lys3}_{UUU}-t⁶A₃₇. Thus, introduction of t⁶A₃₇ did not affect protonation of A₃₈.

While most signals of the spectra could be assigned using standard protocols, site-specific substitutions of ¹⁵N-labeled nucleosides of the unmodified ASL^{Lys3}_{UUU} were used to confirm or to identify the imino proton resonances of the uridine-rich loop. We were able to unambiguously identify the imino protons of loop residues U₃₄, U₃₅ and U₃₆ (Figure 2A). Comparison of the one-dimensional spectra of the t⁶A₃₇-modified and unmodified hairpin indicated a major rearrangement of the loop uridines with the introduction of t⁶A₃₇. In the one- and two-dimensional spectra of the unmodified ASL^{Lys3}_{UUU}, a base-paired imino resonance at 13.55 ppm was assigned to U₃₃. This was the only imino proton unaccounted for by standard and ¹⁵N-labeling procedures. Its position in the spectrum was that of a canonical U·A base pair that we assigned to U₃₃·A₃₇ in the unmodified ASL. However, the signal was not observed in spectra of ASL^{Lys3}_{UUU}-t⁶A₃₇. Instead, a fourth free-imino signal from a uridine was observed between 10.9 and 11.4 ppm. This is the expected result from modification of the exocyclic amine of A₃₇ because t⁶A₃₇ would be unable to form a canonical base pair with U₃₃. Thus, the introduction of t⁶A₃₇ negated an intraloop U₃₃·A₃₇ base pair that was observable in the unmodified ASL^{Lys3}_{UUU}. Loss of the intraloop base pair was probably responsible for the lower T_m of ASL^{Lys3}_{UUU}-t⁶A₃₇, relative to that of the unmodified ASL^{Lys3}_{UUU} (Table 1).

Assignment of Nonexchangeable Resonances. Assignment of the nonexchangeable resonances relied on a variety of homo- and heteronuclear experiments. Initial assignment of the aromatic H5-H6 protons relied on assigned nonex-



Figure 2. Dependence of the rate of polymerization (R_p) on the concentration of the initiator (I_0). (A) $R_p \propto I_0$; (B) $R_p \propto I_0^{-1/2}$.

The dependence of the rate of polymerization (R_p) on the concentration of the initiator (I_0) is shown in Figure 2. Curve (A) shows a linear relationship, indicating a first-order dependence of R_p on I_0 . Curve (B) shows a non-linear relationship, indicating a half-order dependence of R_p on I_0 . The transition from first-order to half-order behavior is characteristic of a reaction mechanism involving a pre-equilibrium between the initiator and a reactive species.

The rate of polymerization (R_p) is defined as the rate of change of the concentration of the polymer (P) with respect to time (t):

$$R_p = \frac{dP}{dt}$$

The concentration of the initiator (I_0) is defined as the total concentration of the initiator species at the start of the reaction. The concentration of the initiator (I) at any time (t) is given by:

$$I = I_0 - 2kt_p t$$

where k_t is the rate constant for the termination reaction and t_p is the average lifetime of the polymer radical. The rate of polymerization (R_p) is given by:

$$R_p = k_p [M] [P]$$

where $[M]$ is the concentration of the monomer and $[P]$ is the concentration of the polymer radical. The concentration of the polymer radical ($[P]$) is given by:

$$[P] = \frac{k_i [I_0]}{k_t}$$

where k_i is the rate constant for the initiation reaction. Substituting this expression for $[P]$ into the equation for R_p gives:

$$R_p = k_p [M] \frac{k_i [I_0]}{k_t}$$

which is the equation for curve (A) in Figure 2. The half-order dependence of R_p on I_0 (curve B) is observed when the initiation reaction is reversible and the concentration of the initiator (I_0) is high enough to establish a pre-equilibrium between the initiator and a reactive species. In this case, the rate of polymerization (R_p) is given by:

$$R_p = k_p [M] \frac{k_i [I_0]^{1/2}}{k_t}$$

which is the equation for curve (B) in Figure 2.

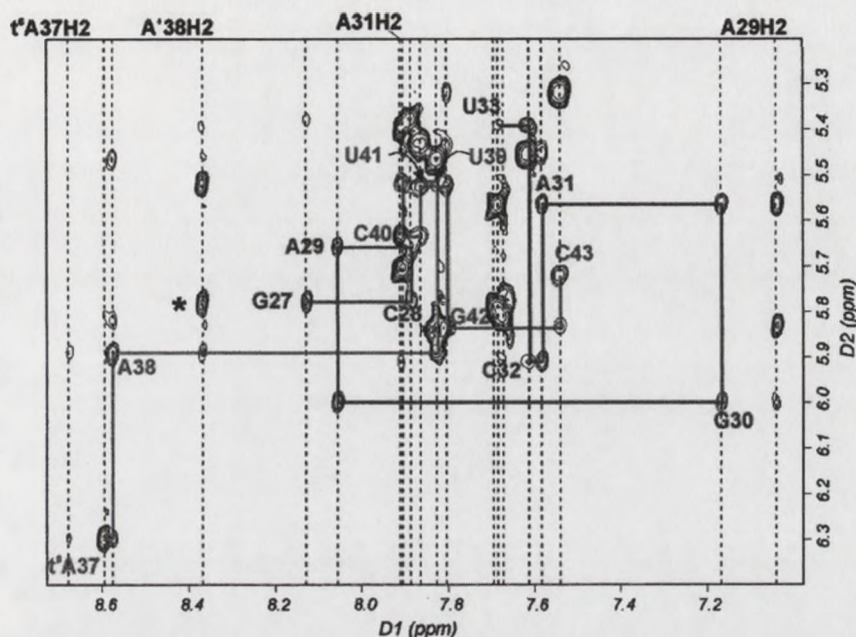


FIGURE 3: NOESY spectrum of ASL^{Lys3}UUU-t⁶A₃₇ in D₂O at 25 °C showing connectivity between the aromatic protons and the H1' of the ribose. The 5' side of the loop is shown in red and the 3' side of the loop is in green. Vertical dashed lines indicate each of the aromatic resonances, H8, H6, or H2. One of the important intraloop cross-peaks, A⁺₃₈H2/U₃₄H1' is marked with an asterisk.

changeable resonances and both DQF-COSY and TOCSY experiments. Differentiation of cytidine from the uracil aromatic protons was aided by natural abundance ¹H-¹³C HSQC spectra. Cross-peaks from HSQC spectra were used to identify H1' protons and distinguished adenosine H2 from purine H8 and pyrimidine H6 resonances. Both adenosines in the stem (A₂₉ and A₃₁) had NOE cross-peaks from H2 to the H1' proton of their 3' neighbor and A₂₉ also had an interstrand cross-peak to the H1' of the 5'-adjacent base pair. Comparison to NOESY spectra of the sample in H₂O enabled us to assign the adenosine H2 resonances. Sequential aromatic-H1' connectivities of the NOESY spectrum (in D₂O) could be traced for most of the molecule (Figure 3). As in the stem, A₃₈ and t⁶A₃₇ stacked as indicated by NOEs to the H1' of their 3' neighbor. Several unique cross-peaks were also observed. There was a cross-peak from the t⁶A₃₇-H2 to the methyl of the side chain (H_γ) and an unexpected NOE from A₃₈H2 across the loop to both H1' and H2' of C₃₂. Sequential aromatic-H1' connectivities could be traced from the 5' terminus of the stem to the first residue of the loop, C₃₂ (Figure 3). At that point there was a break in the sequential connectivity that included the invariant U₃₃. Though NOEs were observed between U₃₅ and U₃₆, sequential aromatic to H1' connectivities began again at t⁶A₃₇ and continued to the 3'-terminal C₄₃ (Figure 3). In the longer mixing time NOESY spectrum (400 ms), sequential connectivities between the nucleobase aromatics (H6, H8) were observed. These connectivities were apparent on the 5' side of the stem from G₂₇ to A₃₁ and on the 3' side of the loop from t⁶A₃₇ up to the 3' terminus, with one exception. The chemical shift of C₄₀H6 was too close to its neighbors to be resolved reliably from the diagonal. Similar to the lack of aromatic to H1' connectivities, no sequential aromatic-to-aromatic resonances were observed from residue A₃₁ through t⁶A₃₇.

Identification of three uridine nucleoside spin systems of the anticodon (U₃₄, U₃₅, U₃₆) was hindered by overlap in the spectra and a lack of resolvable intranucleotide aromatic to ribose NOEs. Several long distance NOEs from already identified nucleosides in the loop to anticodon uridines were chosen for distance geometry analysis to determine the most likely uridine spin system assignment. A set of distance geometry calculations was performed with all assigned distance restraints and one interresidue restraint from an assigned resonance to an unassigned uracil. Then 60 structures were generated with three distance restraint sets (e.g., from A₃₈H2 to either U₃₄H1', U₃₅H1', or U₃₆H1'). After steepest decent minimization, the sets of structures were evaluated based on their total energy and magnitude of restraint violations. Assignment of an NOE to a particular anticodon uridine spin system was accomplished by eliminating those possibilities that yielded highly distorted geometries or high-energy structures.

Both DQF-COSY and TOCSY experiments were used to assign H1' and H2' resolvable cross-peaks. H1'-H2' cross-peaks in the DQF-COSY spectrum were observed for loop residues, U₃₃ to t⁶A₃₇, and for C₄₃. The most intense H1'/H2' cross-peaks were residues U₃₃ to U₃₆. A short mixing time (50 ms) NOESY spectrum also aided in determining H2' resonances. The H3' of the riboses were assigned using the above spectra and the ¹H-³¹P HETCOR spectrum. The same experiments were used to determine more than half of the H4' and H5'/5'' protons.

NOE and Dihedral Angle Restraints. A total of 272 distance and dihedral angle restraints were used in the structure calculations (Table 2). Of the 220 NOE-derived restraints, there were 117 intranucleotide and 103 internucleotide distance restraints. Thirteen restraints were derived from exchangeable proton spectra and were added to mimic the hydrogen-bonding pattern of the stem and to limit fraying

Table 2: Structure Determination Statistics

total NOE distance restraints	220
intranucleotide NOEs	117
internucleotide NOEs	90
hydrogen-bonded stem base pairs	13
dihedral angle restraints	52
refinement statistics (10 lowest-energy structures)	
NOE violations (>0.1 Å)	0
dihedral angle violations (>10°)	0
avg pair wise rmsd (all heavy atoms) (Å)	
all residues	2.30 ± 0.67
stem residues (27–31,39–43)	0.76 ± 0.20
loop residues (32–38)	3.07 ± 1.07

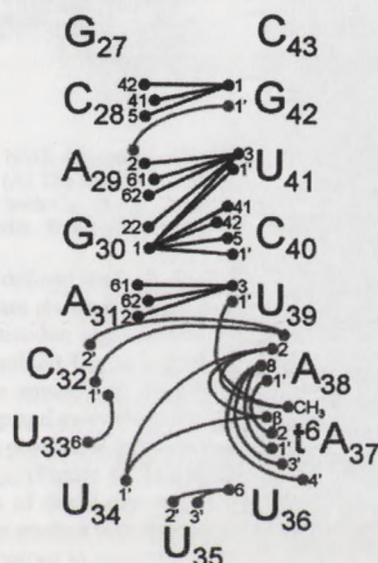


FIGURE 4: Illustration of a portion of the distance restraints used to determine the structure of ASL^{Lys3}_{UUU}-t⁶A₃₇. Restraints derived from exchangeable resonances are depicted in blue while long-range (2 or more residues apart) restraints calculated from nonexchangeable NOE buildup curves are shown in green. Sequential distance restraints between bases are only shown for the loop residues (C₃₂-A₃₈) and are in red.

of the terminal bases (Table 2). In addition, a total of 52 dihedral angle restraints were used for structural calculations. From the DQF-COSY and ³¹P spectra, 46 dihedral angle restraints defining the backbone and sugar pucker were incorporated. The torsion angles between U₃₃ and U₃₄ were left unrestrained to not bias the possibility of a tRNA^{Phe} anticodon-like U-turn, though none was observed in the ³²P-spectra of either the unmodified or t⁶A₃₇-containing ASLs. In addition, five through space dihedral angle restraints were placed between the bases of the stem to maintain their planarity (±10°) and an additional restraint was needed on the side chain of t⁶A₃₇ to keep it planar. The glycosidic angles χ were left unrestrained since the intranucleotide H1'/H6 or H1'/H8 NOE is capable of defining their syn/anti conformation. However, the long-range restraints, some of which are illustrated in Figure 4, were more informative and certainly more important in determining the global fold of the molecule. In particular, the distance restraints between the 5' side of the stem and C₃₂ with A₃₈ and the threonine of t⁶A₃₇ with A₃₈ and U₃₉ of the 3'-side of the stem were critical in defining the stem to loop angle of the ASL. The δ torsion angles for 14 residues were constrained based on the analysis

of a high-resolution DQF-COSY spectrum. Residues U₃₃, U₃₄, and U₃₆ had H1'-H2' coupling constants greater than 7 Hz and therefore were constrained to the C2'-endo range. In contrast, the stem residues had narrow H1'-H2' (<2 Hz) cross-peaks, and were constrained to the C3'-endo range. Where the H3'-H4' cross-peaks could be easily identified, a large coupling constant was observed. Residues with intermediate coupling constants, U₃₅, t⁶A₃₇, and A⁺₃₈, were left unconstrained. On the basis of the ³¹P chemical shifts of all residues, their corresponding α and ζ torsion angles were loosely constrained to exclude the trans conformation, with the exceptions of U₃₃ and U₃₄, which were left unrestrained.

Structural Determination Features. A superposition of the 10 lowest energy structures of the ASL^{Lys3}_{UUU}-t⁶A₃₇ resulted in an average pairwise rmsd of 2.30 ± 0.67 Å for all heavy atoms of the molecule (Figure 5A). The five base paired stem of ASL^{Lys3}_{UUU}-t⁶A₃₇ adopted the form of A-RNA, and all stem nucleosides of the 10 lowest energy structures were C3'-endo. Only U₃₃, U₃₄, and U₃₆ of the loop were found to be strongly C2'-endo. Sixteen glycosidic angles were found to be anti, though no torsion angle restraints were used. Relative to the mean structure, the average rmsd was 1.57 ± 0.33 Å. The stem was much better defined than the loop. The heavy atom average pairwise rmsd of the stem was 0.76 ± 0.20 Å, while that of the loop was 3.07 ± 1.07 Å. Relative to the mean structure, the average rmsd was 0.52 ± 0.14 Å and 2.13 ± 0.52 Å for the stem and loop, respectively. The better rmsd relative to the average structure reflects that there were two somewhat different families of low energy structures. One family had U₃₃ stacked below C₃₂ while the second group of structures had U₃₃ placed adjacent to U₃₄. Similarly in determination of the structure of another seven-membered loop, that of the yeast tRNA^{Phe} T Ψ C loop, we found two almost equally populated families of structures with m⁵U₅₄ (T₅₄) either within the loop or displayed outwardly (42). The H8/H1' NOE between C₃₂ and U₃₃ in ASL^{Lys3}_{UUU}-t⁶A₃₇ is weaker than an H8/H1' internucleotide NOE for A-form RNA. With U₃₃ being the terminal base in the stacking on the 5' side of the stem, weak base stacking could best account for the weaker NOE. For structure comparisons, we used a minimized average of the structures with C₃₂ and U₃₃ stacked (Figure 5B).

RNA hairpins have an angle between the plane of the loop and the axis of the stem that can only be defined by structural restraints between the loop and stem. A number of NOE cross-peaks are derived from C₃₂ on the 5'-side of the loop and t⁶A₃₇ and A₃₈ on the 3'-side of the loop with the adjacent A₃₁-U₃₉ stem base pair. Restraints from these NOEs structure the loop locally at the junction with the stem (Figure 5C) in an angle of 133°, comparable to the 113° of the yeast tRNA^{Phe} X-ray crystallographic structure (52). In addition, RNA loops can be flexible and unstructured, particularly in the middle of the loop and when the loop is as large as that of the anticodon domain. The ASL^{Lys3}_{UUU}-t⁶A₃₇ has a seven-membered loop typical of anticodon domains, but with four uridines in the middle of the sequence. Significant chemical shift similarities of these four tandem uridines (U₃₃, U₃₄, U₃₅, and U₃₆) made structure determination particularly challenging. However, ¹⁵N-labeling and reiterative modeling allowed us to distinguish the uridines spin systems. A total of 69 NOE-derived distance and 16 backbone torsion angle

Table 1. Summary Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Y	1.23	0.45	0.00	3.00
X1	1.56	0.32	0.00	3.00
X2	1.89	0.28	0.00	3.00
X3	2.12	0.25	0.00	3.00
X4	2.35	0.22	0.00	3.00
X5	2.58	0.19	0.00	3.00
X6	2.81	0.16	0.00	3.00
X7	3.04	0.13	0.00	3.00



Figure 1. Heteroscedasticity in the Error Term

The results of the regression analysis are presented in Table 2. The coefficient estimates for the independent variables are all positive and statistically significant at the 5% level. The adjusted R-squared value is 0.85, indicating a strong fit of the model. The F-statistic is 123.45, which is highly significant. The Durbin-Watson statistic is 1.89, suggesting no significant autocorrelation in the residuals. The Breusch-Pagan test results indicate the presence of heteroscedasticity in the error term, with a p-value of 0.001. This is visually confirmed by the scatter plots in Figure 1, which show that the variance of the residuals increases as the predicted values of the independent variables increase.

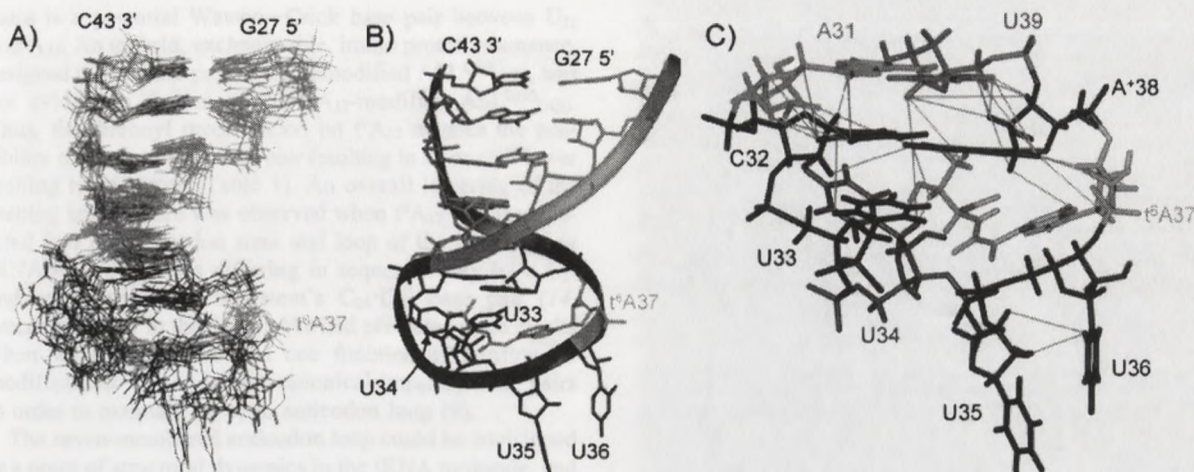


FIGURE 5: NMR derived structure of ASL^{Lys3}-t⁶A₃₇. The 5'-side of the stem is shown in green and 3'-side shown in cyan; t⁶A₃₇ is highlighted in orange. (A) The superposition of 10 lowest energy structures of ASL^{Lys3}_{UUU}-t⁶A₃₇. (B) Minimized average structure of the family where U₃₃ stacks with C₃₂. (C) A close up view of the nucleosides at the base of the stem and the loop of ASL^{Lys3}_{UUU}-t⁶A₃₇ with U₃₄ shown in red for clarity. Hydrogen atoms are displayed with green lines indicating interresidue NOEs.

restraints defined the loop structure and some of the distance restraints are shown in Figure 5C. NOEs limited the positions of the anticodon nucleosides. NOEs from U₃₄ to A⁺₃₈ and t⁶A₃₇ restrained U₃₄ to a position in the loop (Figure 5C). U₃₄ in the unmodified ASL^{Lys3}_{UUU} structure also is located in the loop and away from the other anticodon uridines (15), though its position is different from that in the t⁶A₃₇-modified ASL^{Lys3}_{UUU} (Figure 6). In contrast, Gm₃₄ in the purine-rich anticodon of the X-ray crystallographic structure of yeast tRNA^{Phe} is stacked with the subsequent anticodon bases (52). Also in contrast to yeast tRNA^{Phe}, the lack of any far shifted resonances in the phosphorus spectrum suggests that there is no U-turn found in ASL^{Lys3}_{UUU}-t⁶A in solution. The signature cross-peak of a U-turn, the *n* to *n* + 2 internucleotide H1' to H8 resonance, was not observed due to overlap in the NOESY spectra. Although we did not eliminate the possibility of a U-turn in the dihedral restraints, no structures with U-turns were generated. In fact, U₃₄ could not have NOEs with A⁺₃₈ and t⁶A₃₇ and remain at the bottom of the anticodon loop as in yeast tRNA^{Phe}.

DISCUSSION

Incorporation of t⁶A₃₇ into the otherwise unmodified and nonfunctional ASL^{Lys3}_{UUU} restored ribosome-mediated, poly-A binding (14, preceding paper in this issue). Because the physicochemical contributions of t⁶A₃₇ impart function to the ASL, it was important to determine the ASL^{Lys3}_{UUU}-t⁶A₃₇ structure and define the differences in conformation between it and the unmodified ASL. The structures of the unmodified ASL^{Lys3}_{UUU} and ASL^{Lys3}_{UUU} with Ψ₃₉ in the stem and immediately adjacent to the loop (15) are pertinent to the description of the ASL^{Lys3}_{UUU}-t⁶A₃₇ structure. The presence of pseudouridine demonstrably stabilized the conformation (14, 15). However, we have found that both the unmodified and Ψ₃₉-containing ASL^{Lys3}_{UUU} are nonfunctional, i.e., the ASL constructs will not bind poly-A programmed ribosomes (14, preceding paper in this issue).

The addition of t⁶A₃₇ to the 3'-side of the loop and adjacent to the anticodon of ASL^{Lys3}_{UUU} altered the structure of the molecule in a way that allowed it to bind the ribosome (14,

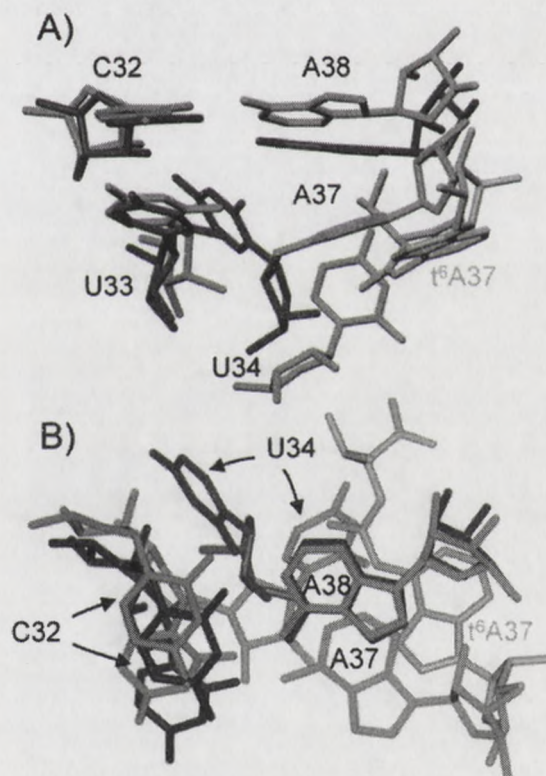


FIGURE 6: Superposition of the ASL^{Lys3}_{UUU}-t⁶A₃₇ (blue) and unmodified ASL^{Lys3}_{UUU}-U₂₇-A₄₃ (green) (15). Phosphate atoms were used to superimpose the structures. The ribose and phosphate backbone along with U₃₅ and U₃₆ are not displayed to improve clarity. (A) Side view showing U₃₄ displacement by the threonyl modification of t⁶A₃₇. (B) Top view showing that the new position of U₃₄ displaces C₃₂.

preceding paper in this issue). Examination of the exchangeable NOESY spectra shows that t⁶A₃₇ does not change the conformation of the stem. Thus, t⁶A's physicochemical contribution to a rearrangement of the loop results in the ASL's AAA codon binding activity. In the unmodified ASL,

The first part of the paper discusses the general principles of the theory of the firm, which are based on the assumption of profit maximization. This leads to the derivation of the firm's supply curve, which is upward sloping in the short run and perfectly elastic in the long run.

The second part of the paper discusses the firm's cost structure, which is determined by the technology of production. The firm's cost function is derived from the production function, and it is shown that the firm's cost function is convex in the short run and linear in the long run. This leads to the derivation of the firm's short-run supply curve, which is upward sloping and perfectly elastic in the long run.

The third part of the paper discusses the firm's long-run supply curve, which is perfectly elastic. This is because the firm can enter or exit the industry at will, and the industry supply curve is perfectly elastic. This leads to the derivation of the long-run equilibrium price, which is equal to the firm's long-run marginal cost.

there is a potential Watson–Crick base pair between U₃₃ and A₃₇. An upfield, exchangeable, imino proton resonance, assigned to this base pair in the unmodified ASL^{Lys3}_{UUU}, was not evident in spectra of the t⁶A₃₇-modified ASL^{Lys3}_{UUU}. Thus, the threonyl modification on t⁶A₃₇ negates the possibility of this intraloop base pair resulting in an overall lower melting temperature (Table 1). An overall lowering of the melting temperature was observed when t⁶A₃₇ was incorporated into the anticodon stem and loop of the isoaccepting tRNA^{Lys1,2}_{CUU} species differing in sequence only by a C₃₄ and an inversion of the stem's C₂₈·G₄₂ base pair (14, preceding paper in this issue). We had predicted these results when we hypothesized that one function of position 37 modifications was to inhibit canonical intraloop base pairs in order to maintain an open anticodon loop (9).

The seven-membered anticodon loop could be envisioned as a point of structural dynamics in the tRNA molecule, and a region in which stacking interactions would not be found. However, hydrogen bonding between t⁶A₃₇ H11 and N1 of the threonyl side chain of t⁶A₃₇ completes an additional planar ring on adenosine and consequently makes it similar to the guanosine-derived, tricyclic base wybutosine-37 of yeast tRNA^{Phe}. The tricyclic nature of t⁶A possibly enhances the base stacking ability of the nucleoside. In the solution structure of the ASL^{Lys3}_{UUU}, t⁶A₃₇ and A⁺₃₈ are coplanar and stacked but because the threonyl side chain must be accommodated in the loop, the two bases are offset by an angle of 40.1°. Stacking interactions are very important in stabilizing RNA structures. Calculations of the entropic thermodynamic parameter (ΔS) for the t⁶A₃₇-modified and unmodified ASL^{Lys3}_{UUU} (Table 1), and for the anticodon stem and loop of the isoaccepting tRNA^{Lys1,2}_{CUU} species (14, preceding paper in this issue), indicate a considerable contribution by t⁶A₃₇. In all structures of all calculations, t⁶A₃₇ is stacked on A₃₈, which is in turn stacked on the adjacent U₃₉ in the stem. Thus, modification at position 37 structures the 3'-side of the loop with stacking to A₃₈ and the adjacent stem.

t⁶A₃₇ displaces U₃₄ relative to its position in the unmodified ASL. However, U₃₄ in proximity to A⁺₃₈ and t⁶A₃₇, this leaves only two residues, U₃₅ and U₃₆ available for hydrogen bonding to a message on the ribosome (Figure 5). With only two of three anticodon bases available for codon pairing, this unconventional anticodon structure is a reasonable explanation for the bacterial and mammalian tRNA^{Lys}_{SUU} tendency to frame shift, prematurely terminate translation and mis-read Asn codons (10). We postulated that the modified nucleoside distortion of the anticodon loop was a possible structural determinant for the preferential selection of tRNA^{Lys3}_{SUU} as primer of HIV-1 reverse transcriptase *in vivo* (10). U₃₅ and U₃₆, though not well defined in the structure for lack of NOE restraints to other nucleosides, were displayed outward from the 10 lowest energy structures (Figure 5A). NOEs between the ribose of U₃₅ and the H6 of U₃₆ helped define their positions relative to each other. Their placement is consistent with that found in the X-ray crystallographic structure of lysyl-tRNA synthetase with cognate tRNA^{Lys} wherein a phenylalanine intercalates between U₃₅ and U₃₆ (53).

At the pH 5.6, A⁺₃₈ is protonated and charged in the unmodified ASL^{Lys3}_{UUU} (15). Incorporation of t⁶A₃₇ does not alter the protonation of A⁺₃₈. However, the negative charge contributed by the dissociated carboxylic acid of threonine

would neutralize the positive charge of A⁺₃₈. While the H1 proton of A⁺₃₈ is not directly observable, protonation of the N1 of the nucleobase is evident from the chemical shift change in H2 (15, 54). An even clearer indicator of adenosine protonation is the upfield shift of the C2 resonance (54). As A₃₈ became protonated with the lowering of the pH, its C2 resonance shifted upfield (see Supporting Information). Formation of an intraloop base pair between C₃₂ and A⁺₃₈ and adjacent to the stem, could be responsible for stabilizing the RNA at low pH (Table 1) (15). Introduction of t⁶A₃₇ altered the C₃₂·A⁺₃₈ base-pairing geometry found in the unmodified ASL^{Lys3}_{UUU} (15). The threonyl group occupies the same space as U₃₄ of the unmodified ASL^{Lys3}_{UUU} (Figure 6). To adapt to the t⁶A modification, U₃₄ is displaced away from t⁶A₃₇ and C₃₂ is moved to accommodate U₃₄ (Figure 6). Displacement of U₃₄ by t⁶A₃₇ distorts the C₃₂·A⁺₃₈ base pair by shifting C₃₂ out of alignment with A⁺₃₈. The through space hydrogen bonding angles, A₃₈–N1–H1 to C₃₂–O2 and A₃₈–N6–H6.1 to C₃₂–N3 are 163° and 168°, respectively, for the unmodified ASL^{Lys3}_{UUU}. When t⁶A₃₇ is added, the hydrogen bonding angles decrease to 143° and 155°, an average of a 17° difference from the unmodified structure and 31° from being linear. It should be noted that this geometry of structure arises from an analysis of the average structure, and that no hydrogen-bonding constraints involving A₃₈ were introduced in the calculations. When C₃₂·A⁺₃₈ was constrained to be hydrogen-bonded and repeated with the same restrain set, no new violations arose during the distance geometry phase of the modeling and after refinement there was little difference in energy distribution among the structures (data not shown).

Most significant to biological function, t⁶A₃₇ eliminates the canonical intraloop U₃₃·A₃₇ hydrogen bond, displaces U₃₄ and facilitates 3'-stacking of the loop. Because t⁶A₃₇ negates base pair formation in the ASL^{Lys3}_{UUU} loop no additional rearrangement of the bases need occur to accommodate the modifications of U₃₄, mcm⁵s²U₃₄, and the ms²-derivative of t⁶A₃₇ (Figures 5 and 6). From these results and those of the preceding paper in this issue (14), we conclude that an open, structured loop is required for ribosomal binding. Ribosome binding cannot simply be an induced fit of the ASL, otherwise the unmodified ASL^{Lys3}_{UUU} would bind as well as the unmodified ASL^{Lys1,2}_{CUU} (14, preceding paper in this issue). This not being the case, anticodon stem and loop domains are distinctive for unambiguous recognition of the tRNA by cognate aminoacyl-tRNA synthetase and for recognition of individual codons on the ribosome, yet share a similar architecture for ribosome binding. Modified nucleoside chemistry and structure impart just such a duality to ASL^{Lys3}_{UUU} and probably function similarly for other tRNAs.

ACKNOWLEDGMENT

A special thanks to Robert J. Cain whose initial studies demonstrated that this project was feasible. We thank Winnell Newman and Guihua Liu of the NCSU Nucleic Acids Facility for the synthesis and purification of the RNAs used in this study. We would also like to thank Robert J. Cain and Connie Yarian for editorial comments and Hanna Sierzputowska-Gracz of the NCSU NMR Facility for help in spectroscopy.



SUPPORTING INFORMATION AVAILABLE

Graph summarizing the translational diffusion experiments in D₂O at 25 °C, high-resolution DQF-COSY and HSQC spectra of ASL^{Lys3-t6}A₃₇, and a superposition of three HSQC spectra highlighting the pH dependent shift of A₃₈C₂. Coordinates of the 10 lowest energy structures and an average structure have been deposited in the Protein Data Bank with accession number 1feq. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

1. Tsuchihashi, Z., and Brown, P. O. (1992) *Genes Dev.* 6, 511–9.
2. Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J., and Varmus, H. E. (1988) *Nature* 331, 280–3.
3. Precup, J., and Parker, J. (1987) *J. Biol. Chem.* 262, 11351–5.
4. Hagervall, T. G., Pomerantz, S. C., and McCloskey, J. A. (1998) *J. Mol. Biol.* 284, 33–42.
5. Heurgue-Hamard, V., Mora, L., Guarneros, G., and Buckingham, R. H. (1996) *EMBO J.* 15, 2826–33.
6. Isel, C., Marquet, R., Keith, G., Ehresmann, C., and Ehresmann, B. (1993) *J. Biol. Chem.* 268, 25269–72.
7. Isel, C., Lanchy, J. M., Le Grice, S. F. J., Ehresmann, C., Ehresmann, B., and Marquet, R. (1996) *EMBO J.* 15, 917–24.
8. Auxilien, S., Keith, G., Le Grice, S. F., and Darlix, J. L. (1999) *J. Biol. Chem.* 274, 4412–20.
9. Dao, V., Guenther, G., Malkiewicz, A., Nawrot, B., Sochacka, E., Kraszewski, A., Everett, K., and Agris, P. F. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 2125–9.
10. Agris, P. F., Guenther, R., Ingram, P. C., Basti, M. M., Stuart, J. W., Sochacka, E., and Malkiewicz, A. (1997) *RNA* 3, 420–8.
11. Agris, P. F., Söll, D., and Seno, T. (1973) *Biochemistry* 12, 4331–7.
12. von Ahsen, U., Green, R., Schroeder, R., and Noller, H. F. (1997) *RNA* 3, 49–56.
13. Ashraf, S. S., Sochacka, E., Cain, R., Guenther, R., Malkiewicz, A., and Agris, P. F. (1999) *RNA* 5, 188–94.
14. Yarian, C., Marszalek, M., Sochacka, E., Malkiewicz, A., Guenther, R., Miskiewicz, A., and Agris, P. F. (2000) *Biochemistry* 39, 13390–13395.
15. Durant, P. C., and Davis, D. R. (1999) *J. Mol. Biol.* 285, 115–31.
16. Ogilvie, K. K., Usman, N., Nicoghossian, K., and Cedergren, R. J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 5764–8.
17. Agris, P. F., Malkiewicz, A., Kraszewski, A., Everett, K., Nawrot, B., Sochacka, E., Jankowska, J., and Guenther, R. (1995) *Biochimie* 77, 125–34.
18. Gehrke, C. W., Kuo, K. C., McCune, R. A., Gerhardt, K. O., and Agris, P. F. (1982) *J. Chromatogr.* 230, 297–308.
19. Yarian, C. S., Basti, M. M., Cain, R. J., Ansari, G., Guenther, R. H., Sochacka, E., Czerwinska, G., Malkiewicz, A., and Agris, P. F. (1999) *Nucleic Acids Res.* 27, 3543–9.
20. Serra, M., and Turner, D. H. (1995) *Methods Enzymol.* 249, 242–61.
21. Stejskal, E. O., and Tanner, J. E. (1965) *J. Chem. Phys.* 42, 288–92.
22. Tanner, J. E. (1970) *J. Chem. Phys.* 52, 2523–6.
23. Lapham, J., Rife, J. P., Moore, P. B., and Crothers, D. M. (1997) *J. Biomol. NMR* 10, 255–62.
24. Yang, X., Sanghvi, Y. S., and Gao, X. (1997) *J. Biomol. NMR* 10, 383–8.
25. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* 2, 661–5.
26. Kumar, A., Ernst, R. R., and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun.* 95, 1–6.
27. Macura, S., and Ernst, R. R. (1980) *Mol. Phys.* 41, 95–117.
28. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson.* 65, 355–60.
29. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. R. (1988) *J. Am. Chem. Soc.* 110, 7870–2.
30. Piantini, U., Sorensen, O. W., and Ernst, R. R. (1982) *J. Am. Chem. Soc.* 104, 6800–1.
31. Braunschweiler, L., Bodenhausen, G., and Ernst, R. R. (1983) *Mol. Phys.* 48, 535–60.
32. Mearns, T. H., and Freeman, R. (1983) *J. Magn. Reson.* 51, 531–5.
33. Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T., and Bax, A. (1986) *FEBS Lett.* 208, 94–8.
34. Kellogg, G. W., Szewczak, A. A., and Moore, P. B. (1992) *J. Am. Chem. Soc.* 114, 2727–8.
35. Varani, G., and Tinoco, I., Jr. (1991) *J. Am. Chem. Soc.* 113, 9349–54.
36. Palmer, A. G., III, Cavanagh, J., Wright, P. E., and Rance, J. (1991) *J. Magn. Reson.* 93, 151–70.
37. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc.* 114, 10663–5.
38. Guntert, P., and Wüthrich, K. (1992) *J. Magn. Reson.* 96, 403–7.
39. Wüthrich, K., Billeter, M., and Braun, W. (1983) *J. Mol. Biol.* 169, 949–61.
40. Varani, G., Aboul-ela, F., and Allain, F. H. T. (1996) *Prog. Nucl. Magn. Reson. Spectrosc.* 29, 51–127.
41. Gorenstein, D. G., (1984) *Phosphorus-31 NMR: Principles and Applications*, Academic Press, Florida.
42. Koshlap, K. M., Guenther, R., Sochacka, E., Malkiewicz, A., and Agris, P. F. (1999) *Biochemistry* 38, 8647–56.
43. Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) *J. Comput. Chem.* 7, 230–53.
44. Reddy, P. R., Hamill, W. D., Jr., Chheda, G. B., and Schweizer, M. (1981) *Biochemistry* 20, 4979–86.
45. Reddy, P. R., Schweizer, M. P., and Chheda, G. B. (1979) *FEBS Lett.* 106, 63–6.
46. Várnagy, K., Jezowska-Bojczuk, M., Swiatek, J., Kozłowski, H., Sóvágó, I., and Adamiak, R. W. (1990) *J. Inorg. Biochem.* 40, 357–63.
47. Dewar, M. J. S., Bingham, R. C., and Lo, D. H. (1977) *J. Am. Chem. Soc.* 99, 4899.
48. Besler, B. H., Merz, K. M., and Kollman, P. A. (1990) *J. Comput. Chem.* 11, 431–9.
49. Parthasarathy, R., Ohrt, J. M., and Chheda, G. B. (1977) *Biochemistry* 16, 4999–5008.
50. Tewari, R. (1995) *Chem. Phys. Lett.* 238, 365–70.
51. Wijmenga, S. S., and van Buuren, B. N. M. (1998) *Prog. Nucl. Magn. Reson. Spectrosc.* 32, 287–387.
52. Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H., Seeman, N. C., and Rich, A. (1974) *Science* 185, 435–40.
53. Cusack, S., Yaremchuk, A., and Tuskalo, M. (1996) *EMBO J.* 15, 6321–34.
54. Sierzputowska-Graczyk, H., Gopal, H. D., and Agris, P. F. (1986) *Nucleic Acids Res.* 14, 7783–801.

BI0013039

