

RĪGAS TEHNISKĀ UNIVERSITĀTE

RIGA TECHNICAL UNIVERSITY

Kristaps JAUDZEMS

**PROTEĪNU STRUKTŪRAS, FUNKCIJAS UN LIGANDU
MIJEDARBĪBU PĒTĪJUMI AR KMR**

Promocijas darba kopsavilkums

**NMR STUDIES OF PROTEIN STRUCTURE, FUNCTION
AND LIGAND INTERACTIONS**

Summary of Doctoral Thesis

Rīga 2011

RĪGAS TEHNISKĀ UNIVERSITĀTE
Materiālzinātnes un lietišķās ķīmijas fakultāte
Lietišķās ķīmijas institūts

RIGA TECHNICAL UNIVERSITY
Faculty of Material Science and Applied Chemistry
Institute of Applied Chemistry

Kristaps JAUDZEMS
Doktora studiju programmas „Ķīmija” doktorants

**PROTEĪNU STRUKTŪRAS, FUNKCIJAS UN LIGANDU
MIJEDARBĪBU PĒTĪJUMI AR KMR**

Promocijas darba kopsavilkums

**NMR STUDIES OF PROTEIN STRUCTURE, FUNCTION
AND LIGAND INTERACTIONS**

Summary of Doctoral Thesis

Zinātniskais vadītājs (supervisor)
Dr. habil. chem., profesors (professor)
E.LIEPIŅŠ

Rīga 2011

UDK 577.112(043.2)
Ja 881 p

Jaudzems K. Proteīnu struktūras, funkcijas un
ligandu mijiedarbību pētījumi ar KMR.
Promocijas darba kopsavilkums.-R.:RTU,
2011.-49 lpp.

Iespiests saskaņā ar promocijas padomes "RTU-
P01" 2011.gada 24.februāra lēmumu

Šis darbs izstrādāts ar Valsts pētījumu programmu «Jaunas zāles un biokorekcijas
līdzekļi: konstruēšana, transportformas un darbības mehānisms» un «Jaunu
profilakses, ārstniecības, diagnostikas līdzekļu un metožu, biomedicīnas
tehnoloģiju izstrāde sabiedrības veselības uzlabošanai» (Nr. 2010. 10-4/VPP-4)
atbalstu.



Šis darbs izstrādāts ar Eiropas Sociālā fonda atbalstu projektā «Atbalsts RTU
doktora studiju īstenošanai».

**PROMOCIJAS DARBS
IZVIRZĪTS ĶĪMIJAS DOKTORA GRĀDA IEGŪŠANAI
RĪGAS TEHNISKAJĀ UNIVERSITĀTĒ**

Promocijas darbs dabaszinātņu ķīmijas doktora grāda iegūšanai tiek publiski aizstāvēts 2011.g. Rīgas Tehniskās universitātes Materiālzinātnes un lietišķās ķīmijas fakultātē, Āzenes ielā 14/24, auditorijā.

OFICIĀLIE RECENZENTI

Profesors, Dr. habil. chem. Raimonds Valters
Rīgas Tehniskā universitāte

Profesors, Dr. habil. sc. nat. Martins Billeter
Gēteborgas Universitāte, Zviedrija

Vadošais pētnieks, Dr. Biol. Kaspars Tārs
Latvijas Biomedicīnas pētījumu un studiju centrs

APSTIPRINĀJUMS

Apstiprinu, ka esmu izstrādājis doto promocijas darbu, kas iesniegts izskatīšanai Rīgas Tehniskajā universitātē dabaszinātņu doktora grāda iegūšanai. Promocijas darbs nav iesniegts nevienā citā universitātē zinātniskā grāda iegūšanai.

Kristaps Jaudzems(Paraksts)

Datums:

Promocijas darbs sagatavots kā vienotu zinātnisko publikāciju kopa, satur ievadu un 5 publikācijas, kas uzrakstītas angļu valodā, kopā 60 lappuses.

PROMOCIJAS DARBA VISPĀRĒJS RAKSTUROJUMS

Tēmas aktualitāte

Nesenais progress genomu sekvenēšanā radījis milzīgu apjomu genomiskās informācijas, kuru nepieciešams funkcionāli interpretēt. Gēnu produktu – proteīnu – telpiskās struktūras sniedz informāciju gan par to bioloģisko funkciju, gan terapeitiskās modulēšanas iespējām. Tas sekmējis jaunas post-genomikas zinātnes, strukturālās genomikas, attīstību, kuras uzdevums ir genoma-mēroga augstas caurlaidspējas proteīnu struktūru noteikšana un funkcionāla anotēšana. Rentgenstaru kristalogrāfija un kodolu magnētiskā rezonanse (KMR) ir divas galvenās metodes, kuras izmanto strukturālās genomikas kontekstā proteīnu struktūru noteikšanai atomu izšķirtspējas līmenī. Divas no promocijas darbā aprakstītajām proteīnu struktūrām noteiktas darba autoram piedaloties šādā strukturālās genomikas projektā "Protein Structure Initiative", Skripta pētniecības institūtā ASV, Nobela prēmijas laureāta prof. K. Vītriha vadībā.

Sakarā ar vispārēju biomedicīnisko pētījumu izmaksu pieaugumu, pēdējā laikā nozarē arvien vairāk izmanto racionālās pieejas dažādu medicīnas, bioloģijas un bioķīmijas problēmu risināšanā. Īpaši augsti tiek vērtēts nozīmīgais mērķproteīnu strukturālās informācijas devums racionālā jaunu zāļu vielu dizainā un terapeitisko līdzekļu izstrādē. KMR spektroskopija ir galvenā fizikālā metode, kuru izmanto proteīnu struktūras un dinamikas pētījumiem to dabiskajā vidē – ūdens šķīdumā. Bez tam, KMR metode ir piemērota proteīnu-ligandu (potenciālo zāļu vielu) mijiedarbību pētījumiem un spēj dot informāciju par enerģētiski vājām saistībām, kas nav pieejamas ar citām metodēm, kā arī ļauj noteikt mijiedarbību vietas uz proteīna virsmas.

Izvēlētie pētījumu objekti ir proteīni ar būtisku bioloģisko nozīmi – DnaI proteīns piedalās DNS kopēšanas procesā, R3H domēns atrodams ap 370 proteīnos, kuru konkrētā funkcija ne vienmēr ir zināma, kamēr pats R3H domēns varētu mijiedarboties ar DNS vai RNS, karnitīna acetiltransferāze (CrAT) iesaistīta enerģijas metabolismā, savukārt Nsp7 proteīns ir būtisks smagā akūtā respiratorā sindroma (SARS) koronavīrusa virulences faktors.

Darba mērķi un uzdevumi

Noteikt *de novo* trīsdimensionālo struktūru šķīdumā proteīniem DnaI, NP_247299.1, Nsp7. Ja iespējams, prognozēt to darbības mehānismu molekulārā līmenī. Analizēt Nsp7 proteīna iekšmolekulāros ātruma procesus un dinamiku ūdenī, izvērtēt to ietekmi uz novēroto polimorfismu kristāliskā un šķīduma stāvokļos.

Noteikt mononukleotīda dezoksiguanozīnmonofosfāta (dGMP) saistības vietu uz R3H domēna virsmas, novērtēt dažādu mononukleotīdu saistības enerģētiskos parametrus, izveidot R3H domēna ligandu selektivitātes modeli. Noteikt kompleksa R3H-dGMP telpisko struktūru.

Analizēt CrAT inhibīciju ar mildronātu, izveidot telpiskās struktūras modeli CrAT un mildronāta kompleksam.

Mērķu īstenošanai izvirzīti sekojoši apakšmērķi: klonēt, ekspresēt baktērijās un hromatogrāfiski attīrīt ar ¹³C un ¹⁵N atomiem iezīmētu proteīnu NP_247299.1 miligramu apjomā.

Pētījumu metodika

Proteīnu KMR paraugi trīsdimensionālo struktūru noteikšanai iegūti, izmantojot rekombinanto ekspresiju un hromatogrāfiskās attīrīšanas metodes. Mērķproteīnu kodējošais gēna fragments ieklonēts E.coli baktēriju celma BL21(DE3) šūnās, izmantojot molekulārās bioloģijas metodes. Proteīnu ekspresija veikta M9 minimālajā barotnē, kas satur 1 g/L $^{15}\text{NH}_4\text{Cl}$ un 4 g/L neiezīmētu D-glikozi vai 1 g/L $^{15}\text{NH}_4\text{Cl}$ un 4 g/L [$^{13}\text{C}_6$]-D-glikozi kā vienīgus slāpekļa un oglekļa avotus. Pēc ekspresijas beigām šūnas lizētas ar ultraskaņu (sonikācija) un šķīstošā frakcija sadalīta, izmantojot afinitātes hromatogrāfiju un gēlfiltrāciju. Sadalīšanai izmantots ÄKTA Purifier 10 hromatogrāfs. KMR spektri uzņemti uz Varian Unity Inova 600 MHz, Bruker Avance 600 un 800 MHz spektrometriem. DnaI proteīna struktūras noteikšanai uzņemti un analizēti 3D NOESY- ^{15}N -HSQC, HNHA, 2D NOESY, TOCSY un DQF-COSY KMR spektri. Proteīna NP_247299.1 struktūra noteikta, izmantojot automatizēto projekciju spektroskopiju proteīna galvenās ķēdes atomu ķīmisko nobīžu pierakstīšanai, kā arī 3D NOESY- ^{15}N -HSQC un NOESY- ^{13}C -HSQC spektrus automatizētai sānu ķēžu ķīmisko nobīžu pierakstīšanai un attālumu ierobežojumu noteikšanai. Proteīna Nsp7 struktūras noteikšanai uzņemti un analizēti 3D HNCA, HNCACB, CBCA(CO)NH, HC(C)H-TOCSY, NOESY- ^{15}N -HSQC un NOESY- ^{13}C -HSQC spektri, dinamikas raksturošanai uzņemts heterokodolu 2D $^{15}\text{N}\{^1\text{H}\}$ -NOE spektrs un 2D ^{15}N -HSQC spektru sērija pēc ūdens nomaiņas paraugā ar deiteroūdeni. Polipeptīda amīdgrupu protonu aizsargfaktori noteikti eksperimentālos datus pielīdzinot matemātiskam modelim pašuzrakstītā programmā. Automatizētai proteīna galvenās ķēdes un sānu ķēžu atomu ķīmisko nobīžu pierakstīšanai izmantotas programmas UNIO-MATCH un UNIO-ATNOS/ASCAN. Spektri interaktīvi analizēti izmantojot programmas XEASY un CARA. Struktūras aprēķinātas ar CYANA un UNIO-ATNOS/CANDID programmām. Struktūru minimizēšanai ūdenī izmatotas CNS un OPALp programmas. Struktūru validēšana un analīze veikta ar programmām MOLMOL, PYMOL, PROCHECK u.c.

R3H domēna un CrAT ligandu saistības pētījumos izmantota KMR titrēšana, novērojot signālu ķīmisko nobīžu vai pusplatuma izmaiņas 1D protonu vai 2D ^{15}N -HSQC spektrus. Saistības konstantes aprēķinātas eksperimentālos datus pielīdzinot matemātiskam modelim programmā MS Excel. Mildronāta un CrAT kompleksa telpiskās struktūras modeļa izveidošanai uzņemti piesātinājuma pārnese starpības (saturation transfer difference, STD) KMR spektri un eksperimentālie dati izmantoti datormodelēšanas rezultātu validēšanai. Datormodelēšanai izmantota Molecular Operating Environment programmu pakete.

Zinātniskā novitāte un galvenie rezultāti

- Noteikta iepriekš nezināmā trīsdimensionālā struktūra šķīdumā DnaI proteīna N-gala domēnam, kas uzrāda jaunu cinku-koordinējošu proteīnu uzbūves tipu. Noteiktās ķīmiskās nobīdes un struktūras koordinātas ievietotas attiecīgi Bioloģiskās magnētiskās rezonanses datu bankā (pieejas kods 15926) un Proteīnu datu bankā (pieejas kods 2K7R).
- Noteikta dezoksiguanidīnmononukleotīda saistības vieta R3H domēna telpiskajā struktūrā. Novērtētas mononukleotīdu saistības konstantes, kas ļauj izdarīt secinājumus par proteīna funkciju un iespējamajiem mijiedarbības

partneriem. Identificēti enerģētiskie un strukturālie faktori, kas nosaka ligandu selektivitāti.

- Izveidots modelis antiišēmiskā līdzekļa mildronāta un CrAT mijiedarbībai, kas parāda, ka mildronāta aktivitāte varētu būt daļēji saistīta ar CrAT inhibīciju un sekojošu enerģijas metabolisma ceļu regulāciju.
- Izstrādāts jauns protokols proteīnu rentgenstaru difrakcijas un KMR struktūru analīzei un salīdzināšanai, kas ļauj interpretēt lokāla rakstura uzbūves atšķirības un izdarīt secinājumus par abu tehniku komplementaritāti.
- Noteikta proteīna Nsp7 trīsdimensionālā struktūra šķīdumā pie pH 6,5 un raksturota tā iekšmolekulārā dinamika, kas ļauj izdarīt secinājumus par proteīna funkciju. Parādīts, ka Nsp7 proteīna atlikumi 70–82 atrodas konformacionālā līdzsvarā starp α -spirāles un nestrukturētu formu, kas varētu būt nepieciešama tā fizioloģiski aktīvā kompleksa ar Nsp8 proteīnu iegūšanai. Nsp7 ķīmiskās nobīdes pie pH 6,5 ievietotas Bioloģiskās magnētiskās rezonanses datu bankā zem pieejas koda 15926. 20 KMR konformēru atomu koordinātas, kas attēlo Nsp7 struktūru šķīdumā pie pH 6,5 ievietotas Proteīnu datu bankā zem pieejas koda 2KYS.
- Demonstrēta automatizētās proteīnu struktūras noteikšanas pielietojamība, nosakot proteīna NP_247299.1 telpisko struktūru. NP_247299.1 ķīmiskās nobīdes ievietotas Bioloģiskās magnētiskās rezonanses datu bankā (pieejas kods 16389) un 20 KMR konformēru atomu koordinātas, kas attēlo struktūru šķīdumā ievietotas Proteīnu datu bankā (pieejas kods 2KLA).

Praktiskais pielietojums

Noteiktās trīsdimensionālās proteīnu struktūras un datus par proteīnu-ligandu mijiedarbību var izmantot jaunu terapeitisko līdzekļu dizainā. Eksperimentālās proteīnu struktūras, it īpaši tādas, kas uzrāda jaunu struktūras tipu, izmantojamas arī kā modelis citu proteīnu struktūru modelēšanai uz aminoskābju sekvences līdzības pamata.

Darba struktūra un apjoms

Promocijas darbs sagatavots kā tematiski vienotu zinātnisko publikāciju kopa par KMR spektroskopijas izmantošanu proteīnu pētījumiem. Tas satur 5 publikācijas anonīmi recenzētos un starptautiski pieejamās datubāzēs (Science Citation Index un Chemical Abstracts) iekļautos zinātniskajos izdevumos. Publikācijas uzrakstītas angļu valodā un to kopējais apjoms ir 47 lapaspuses.

Darba aprobācija un publikācijas

Promocijas darba rezultāti izklāstīti un apspriesti 5 publikācijās:

1. Loscha K.V., Jaudzems K., Ioannou C., Su X.-C., Hill F.R., Otting G., Dixon N.E., Liepinsh E. A novel zinc-binding fold in the helicase interaction domain of the *Bacillus subtilis* DnaI helicase loader// *Nucleic Acids Research*. - 2009. - Vol. 37. - p. 2395-2404.
2. Jaudzems K., Zhulyenkov D., Liepinsh E. R3H domain complexes with mononucleotides// *Scientific Journal of RTU. Material Science and Applied Chemistry*. - 2009. - Vol. 20. - p. 58-63.
3. Jaudzems K., Kuka J., Gutsaits A., Zinovjevs K., Kalvinsh I., Liepinsh E., Liepinsh E., Dambrova M. Inhibition of carnitine acetyltransferase by mildronate, a regulator of energy metabolism// *Journal of Enzyme Inhibition and Medicinal Chemistry*. - 2009. - Vol. 24. - p. 1269-1275.
4. Jaudzems K., Geralt M., Serrano P., Mohanty B., Horst R., Pedrini B., Elsliger M.-A., Wilson I.A., Wüthrich K. NMR structure of the protein NP_247299.1: comparison with the crystal structure// *Acta Crystallographica Section F*. - 2010. - Vol. 66. - p. 1367-1380.
5. Johnson M.A., Jaudzems K., Wüthrich K. NMR structure of the SARS coronavirus nonstructural protein Nsp7 in solution at pH 6.5// *Journal of Molecular Biology*. - 2010. - Vol. 402. - p. 619-628.

Darba rezultāti prezentēti sekojošās konferencēs un zinātniskos semināros:

1. Jaudzems K., Zhulyenkov D., Liepinsh E. R3H domain complexes with mononucleotides. Poster at the RTU 49th International Scientific Conference. Riga, October 13, 2008.
2. Jaudzems K. Automation in the NMR structure determination of a 206-residue protein. Oral presentation at the Joint Center for Structural Genomics Annual Meeting. San Diego, USA, April 23, 2009.
3. Jaudzems K., Mohanty B., Serrano P., Geralt M., Liepinsh E., Wüthrich K. Solution NMR structure of the protein NP_247299.1: Comparison with the crystal structure. Poster at the RTU 50th International Scientific Conference. Riga, October 16, 2009.
4. Jaudzems K. Uz KMR balstīta strukturālā genomika ASV programmas "Protein Structure Initiative" ietvaros. Seminārs Latvijas Organiskās sintēzes institūtā. Rīgā, 2010. gada 6. janvārī.
5. Jaudzems K., Geralt M., Serrano P., Mohanty B., Horst R., Pedrini B., Elsliger M.-A., Wilson I.A., Wüthrich K. NMR structure of the protein NP_247299.1: comparison with the crystal structure. Poster at the Joint EUROMAR 2010 and 17th ISMAR Conference. Florence, Italy, July 4-9, 2010.
6. Jaudzems K., Zhulyenkov D., Otting G., Liepinsh E. Structural and thermodynamic characterization of nucleotide binding to the R3H domain from human *Subp-2*. Poster at the 18th International Conference of Methods in Protein Structure Analysis. Uppsala, Sweden, August 25-28, 2010.
7. Jaudzems K. Nucleotide binding to the R3H domain studied by NMR. Oral presentation at the 1st EAST-NMR Young Investigators Meeting. Istanbul, September 29, 2010.

PROMOCIJAS DARBA SATURS

Promocijas darbā iekļauto zinātnisko publikāciju vienojošā tematika ir KMR spektroskopijas izmantošana proteīnu struktūras, funkcijas un ligandu mijiedarbību pētījumiem. Promocijas darba kopsavilkums sadalīts divās nodaļās – pirmajā apkopotas publikācijas, kurās aprakstīta proteīnu telpiskās struktūras noteikšana, bet otrajā – publikācijas, kurās aprakstīti proteīnu mijiedarbību ar ligandiem pētījumi. Katra publikācija atspoguļo atšķirīgu pieeju un metožu izmantošanu dažādu proteīnu vai to ligandu īpašību noteikšanā, kopumā veidojot ganrīz pilnīgu pieejamo proteīnu KMR metožu pielietojuma aprakstu. Nodaļas tālāk sadalītas apakšnodaļās, kas veltītas katras zinātniskās publikācijas kodolīgam izklāstam, iekļaujot īsu ievadu, galvenos rezultātus un secinājumus.

1. NODAĻA. PROTEĪNU STRUKTŪRAS NOTEIKŠANA AR KMR

Proteīnu telpiskās struktūras noteikšana ar KMR ietver vairākus secīgus soļus: mM koncentrācijas proteīna parauga sagatavošanu, KMR spektru uzņemšanu, KMR signālu attiecināšanu individuāliem atomiem polipeptīda molekulā, KMR datu strukturālu interpretāciju un visbeidzot telpiskās struktūras aprēķināšanu un tās enerģijas minimizāciju ūdenī, no kuriem laikietilpīgākie ir pirmie trīs.

KMR eksperimentālās stratēģijas un attiecīgi arī spektru interpretācijas stratēģijas ir tieši atkarīgas no tā, vai ir pieejams neiezīmēts, ^{15}N iezīmēts, vai divkārti ^{13}C un ^{15}N iezīmēts proteīna paraugs. Papildus iezīmes dod iespēju izmantot augstākas dimensijas eksperimentus, kas ievērojami paaugstina signālu izšķirtspēju un atvieglo spektru interpretāciju, kā arī ļauj iegūt kvalitatīvākas telpiskās struktūras. Šajā darbā aprakstītas stratēģijas ^{15}N (DnaI-N piemērs) un divkārti ^{13}C un ^{15}N iezīmētu proteīnu (NP_247299.1 un Nsp7 piemēri) analīzei. NP_247299.1 gadījumā aprakstīta automatizētās projekciju spektroskopijas izmantošana, kas ietver ļoti augstas dimensijas (4D un 5D) eksperimentus, kā arī daļēja spektru interpretācijas automatizēšana laika ekonomijas nolūkos.

Papildus telpiskās struktūras noteikšanai, iespējams veikt eksperimentus, kas raksturo atsevišķu proteīna daļu kustīgumu, strukturālo stabilitāti un konformacionālo līdzsvaru šķīdumā. To izmantošana aprakstīta NP_247299.1 un Nsp7 darbos. NP_247299.1 gadījumā aprakstīts arī jauns protokols kristālisko un šķīduma struktūru salīdzināšanai.

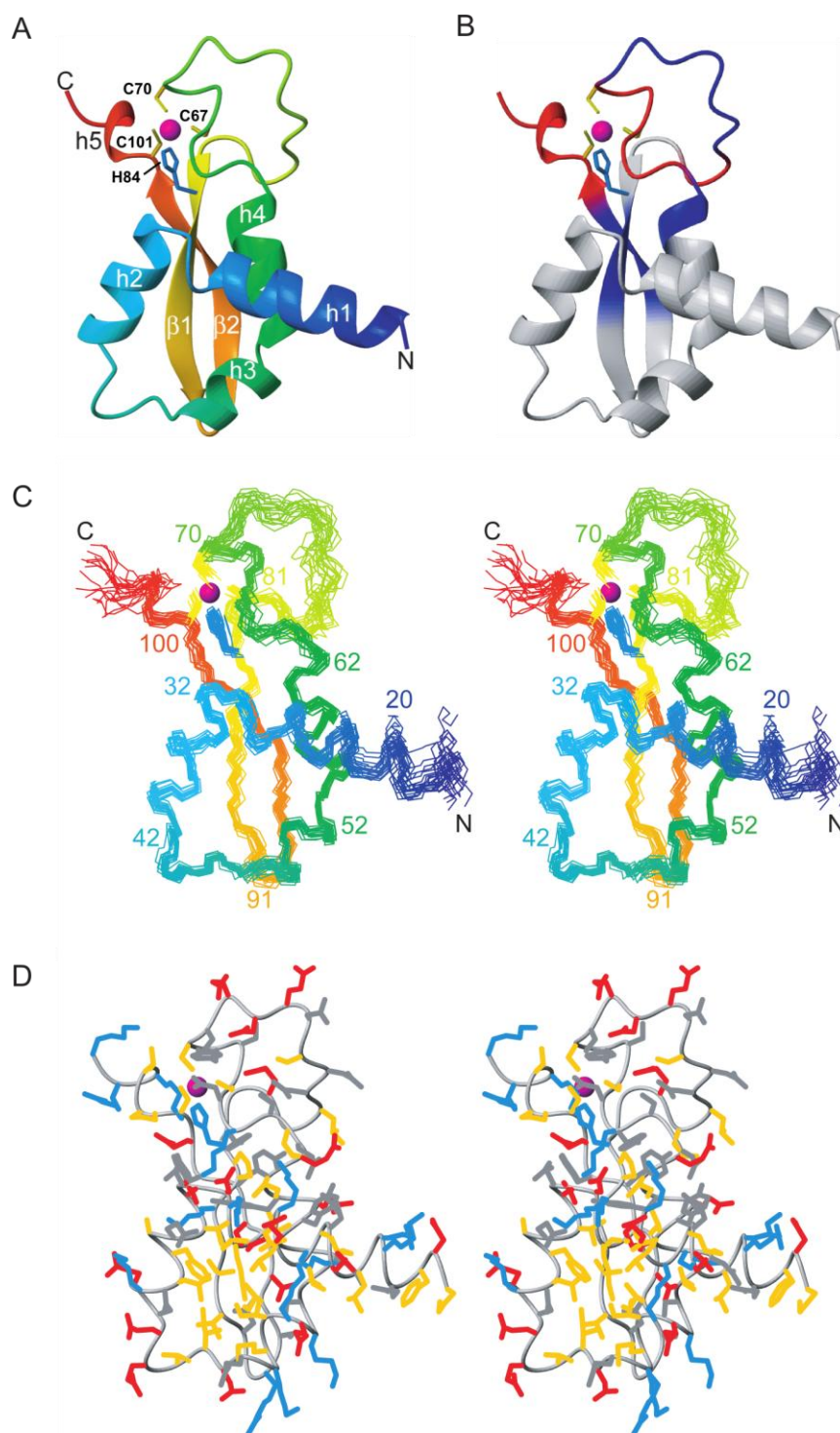
1.1. HELIKĀZES MOBILIZĒTĀJ-PROTEĪNA DOMĒNS DnaI-N UZRĀDA JAUNU CINKU-KOORDINĒJOŠAS STRUKTŪRAS TIPU

DNS replikāciju visos organismos veic replisomas - īpašas multiproteīnu “mašīnas”, kas satur gredzenveida helikāzi DNS dubultspirāles atvīšanai, veidojot divas vienpavediena DNS virknes [1]. Baktērijās, replisomas veidošanās sākumstadijā, helikāze, mijiedarbojoties ar tās mobilizētāj-proteīniem, veido heksamēru struktūru ap vienpavediena DNS. *Bacillus subtilis* organismā helikāzes mobilizācija uz vienpavediena DNS notiek trīs proteīnu – DnaI, DnaB un DnaD – darbības rezultātā [2], bet *in vitro* parādīts, ka DnaI viens pats ir pietiekams helikāzes mobilizācijai [3-4]. DnaI satur divus strukturētus domēnus – N-gala domēnu, kas atbildīgs par mijiedarbībām ar helikāzi un C-gala domēnu, kas saista vienpavediena DNS [4]. Šajā

darbā ar kodolmagnētiskās rezonanses spektroskopiju pirmo reizi noteikta DnaI N-gala helikāzes mijiedarbības domēna (DnaI-N) trīsdimensionālā struktūra un identificēti cinku-koordinējošie aminoskābju atlikumi.

Izmantojot KMR spektroskopiju nevar tieši noteikt cinka jona saistības partnerus uz proteīna, tāpēc cinka ligandu identifikācijai darbā izmantotas divas neatkarīgas metodes. DnaI-N satur vienu histidīnu un četrus cisteīnus, kuri varētu piedalīties koordinēšanā [4]. His84 atlikuma dalību cinka koordinēšanā identificē, uzņemot ^{15}N -HSQC spektru ar INEPT laika periodu 20 ms, kas atbilst koherencei caur divām saitēm $^2J_{\text{HN}}$. His84 Nε2 rezonanses ^{15}N ķīmiskā nobīde uzrādīja koordinētam imidazola slāpeklim raksturīgu vērtību (221 ppm) [5], apstiprinot His84 sānu ķēdes dalību cinka koordinēšanā. Turklāt, KMR spektri neuzrādīja nekādu neviendabīgumu, norādot uz cinka klātbūtni visās proteīna molekulās un identisku koordinācijas vidi. Cinka-koordinējošo cisteīnu identifikācijai izmanto tiolu titrēšanas metodi, kas dabiskā proteīna gadījumā uzrādīja divfāzu pseido-pirmās-pakāpes kinētiku, norādot gan uz koordinētu, gan nekoordinētu cisteīnu klātbūtni. Savukārt, Cys76Ala mutanta gadījumā novēro vienfāzes kinētiku, identificējot Cys76 kā vienīgo cisteīnu, kas nav iesaistīts cinka koordinācijā.

Noteiktā DnaI-N struktūra satur četras α -spirāles, divas β -virknes, kā arī īsu 3_{10} spirāli proteīna C-gala tuvumā (1. att. A). Strukturāli līdzīgu proteīnu meklējumi Proteīnu datu bankā, izmantojot programmu Dali [6], bija nesekmīgi, norādot, ka DnaI-N struktūra uzrāda jaunu telpiskās uzbūves tipu. Cinka jonu koordinē Cys67, Cys70, His84 un Cys101 atlikumi un tam visdrīzāk ir strukturāli nozīmīga loma, jo tas satur kopā trīs aminoskābju sekvencē attālus polipeptīda segmentus ar mazu hidrofobo aminoskābju īpatsvaru (1. att. D). Pirmās 14 N-gala aminoskābes regulāru struktūru neveido. To KMR signāli bija ievērojami šaurāki, norādot uz paaugstinātu kustīgumu. Starp šiem atlikumiem ir vairākas hidrofobās aminoskābes, kas saglabājas visās baktēriju sugās, liecinot par to funkcionālu lomu, iespējams mijiedarbībā ar helikāzi. Tāpat atlikumu 63–82 signāli bija sašaurināti un mazintensīvi. Pagarinātas DnaI konstrukcijas DnaI-N123 gadījumā šajā rajonā novērojām ievērojamas galvenās ķēdes amīdgrupu ķīmisko nobīžu izmaiņas (1. att. B), kas varētu liecināt, ka pagarinātā domēnā šī cilpa, mijiedarbojoties ar DnaI-N C-galam sekojošo segmentu, ieņem noteiktu, nekustīgu konformāciju. Šo pašu secinājumu iegūstam, novērtējot T_2 relaksācijas laika konstantes pagarinātās konstrukcijas atlikumiem 107–123, kuru signāli ir ievērojami paplašināti, bet ķīmiskās nobīdes atbilst nestrukturētai ķēdei.



1. att. DnaI-N struktūra šķīdumā, stereoskati. Attēloti tikai regulāru struktūru veidojošie aminoskābju atlikumi 15–106. Cinka atoms attēlots kā lode (fuksīns) un cinku koordinējošie atlikumi Cys67, Cys70, His84 un Cys101 attēloti ar nūjiņām.

(A) Lentīšu ("ribbon") attēlojums ar otrējās struktūras elementu identifikāciju.

(B) Tas pats, kas (A), bet iekrāsots atspoguļojot galvenās ķēdes amīdgrupu ķīmisko nobīžu izmaiņas DnaI-N pret DnaI-N123 (zils: nelielas izmaiņas $>0,015$ ppm (^1H) vai $>0,15$ ppm (^{15}N); sarkans: ievērojamas izmaiņas $>0,05$ ppm (^1H) vai $>0,5$ ppm (^{15}N)).

(C) 20 KMR konformēru galvenās ķēdes atomu superpozīcija.

(D) Smago atomu attēlojums ar nūjiņām, stereoskats. Sānu ķēdes iekrāsotas zilas (Lys, Arg, His), sarkanas (Asp, Glu), dzeltenas (Ala, Cys, Ile, Leu, Met, Phe, Pro, Trp, Val) un pelēkas (Asn, Gln, Ser, Thr, Tyr)

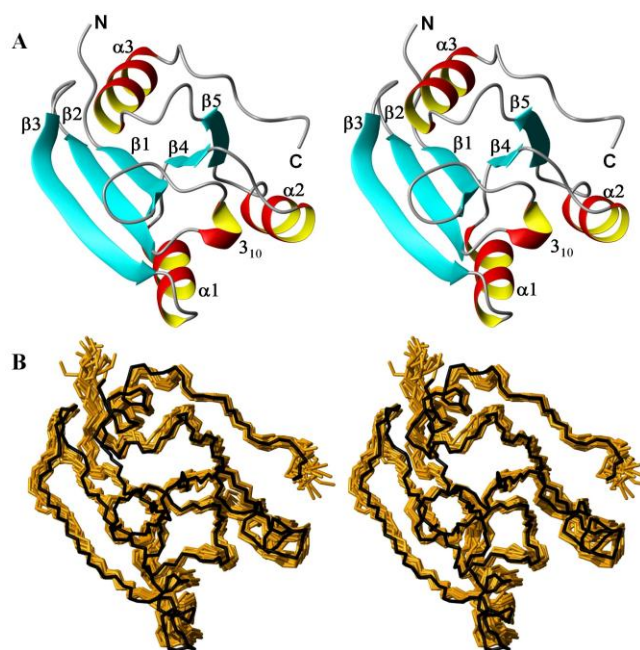
1.2. PROTEĪNA NP_247299.1 KMR STRUKTŪRA ŠĶĪDUMĀ UN SALĪDZINĀJUMS ĀR STRUKTŪRU KRISTĀLISKĀ STĀVOKLĪ

Methanococcus jannaschii genoma lokusa MJ0327 produkta – hipotētiskā proteīna NP_247299.1 trīsdimensionālā struktūra noteikta ASV strukturālās genomikas projekta "Protein Structure Initiative" ietvaros, ar mērķi novērtēt automatizētās proteīnu KMR struktūras noteikšanas protokola pielietojamību un rezultātus. Šajā darbā aprakstīta NP_247299.1 struktūras noteikšana šķīdumā, izmantojot jaunu, automatizētu KMR protokolu, kā arī veikts sistemātisks salīdzinājums ar iepriekš noteiktu struktūru kristāliskā stāvoklī (PDB kods 2qtd).

Proteīna galvenās ķēdes attiecināšanai izmantojām automatizēto projekciju spektroskopiju (APSY-NMR; [7-8]) un UNIO-MATCH programmu [9], kas deva $^1\text{H}^{\text{N}}$, ^{15}N , $^{13}\text{C}^{\alpha}$ un $^1\text{H}^{\alpha}$ atomu ķīmiskās nobīdes 97 (94%) un $^{13}\text{C}^{\beta}$ nobīdes 91 no 104 aminoskābju atlikumiem. Pēc interaktīvas validācijas un neattiecināto atlikumu ķīmisko nobīžu pierakstīšanas veicām automatizētu sānu ķēžu attiecināšanu ar programmu UNIO-ATNOS/ASCAN [10]. Pārbaudot iegūtās sānu ķēžu ķīmiskās nobīdes secinājām, ka aptuveni 95% ir noteiktas pareizi, kā arī iespēju robežās paplašinājām to interpretāciju. Struktūras aprēķināšanai izmantojām programmas UNIO-ATNOS/CANDID [11-12] un CYANA [13], kas veic automatizētu pīķu interpretāciju un attāluma ierobežojumu noteikšanu. Kopsummā interaktīvo soļu veikšana prasīja aptuveni 70 h, kas ir ievērojami mazāk laika nekā gadījumā, ja struktūras noteikšanai netiktu izmantotas automatizētās metodes (tad tam varētu būt nepieciešami vairāki mēneši).

Noteiktā NP_247299.1 KMR struktūra satur trīs α -spirāles, piecas β -virknes un īsu 3_{10} -spirāli, kas sakārtotas secībā $\beta 1-3_{10}-\beta 2-\beta 3-\alpha 1-\beta 4-\alpha 2-\beta 5-\alpha 3$ (2.att. A). β -virknes veido izliektu β -lapu ar topoloģiju 3-2-1-4-5, kurā virknes $\beta 3$, $\beta 1$, $\beta 4$ un $\beta 5$ ir savstarpēji paralēlas, bet $\beta 2$ antiparalēla. Spirāles $\alpha 1$ un $\alpha 2$ novietotas paralēli zem neliela leņķa, ar 3_{10} -spirāli starp to N-galiem. Šīs trīs spirāles pieguļ vienai β -lapas pusei, bet spirāle $\alpha 3$ atrodas pretējā pusē. Pēc superpozīcijas ar kristāliskā stāvokļa struktūru (2. att. B) redzama augsta abu struktūru līdzība.

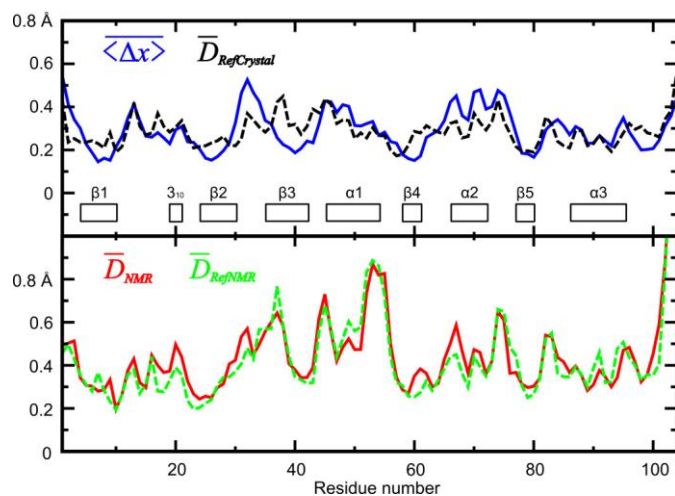
Lai izvērtētu automatizētā protokola rezultātus, veicām detalizētu salīdzināšanu ar struktūru kristāliskā stāvoklī. Strukturāli atšķirīgu proteīna daļu identifikācijai salīdzinājām kristalogrāfiskos B-faktorus ar atomu pozicionālo novirzi starp 20 KMR konformēriem (3. att), kā arī proteīna galvenās ķēdes un sānu ķēžu dihedrālo leņķu vērtības un atomu pakošanos. Identificēto "karsto punktu" precīzākai raksturošanai šķīdumā uzņēmām papildus KMR spektrus, kuros atšķirīgi no $^1\text{H}-^1\text{H}$ kodolu Overhauzera efektiem aproksimēts proteīna konformāciju dinamiskais ansamblis, proti, noteicām skalārās $^1\text{H}^{\text{N}}-^1\text{H}^{\alpha}$ spinu sadarbības konstantes, $^3J_{\text{HN}\alpha}$. No salīdzinājuma secinājām, ka struktūras kodolu veidojošie aminoskābju atlikumi, kam šķīdinātāja piekļuve ir zemāka par 15%, novietojas praktiski identiskā konformācijā. Vietēja rakstura atšķirības novērojamas segmentiem ar augstiem kristalogrāfiskajiem B faktoriem, kas korelē ar dinamisku struktūrālu polimorfismu šķīdumā, piemēram, struktūra šķīdumā uzrāda lēnu dinamisku līdzsvaru starp C-gala prolīna atlikuma cis- un trans-formu, kas pārstāv attiecīgi 25% un 75% no kopējā proteīna šķīdumā (4. att.).



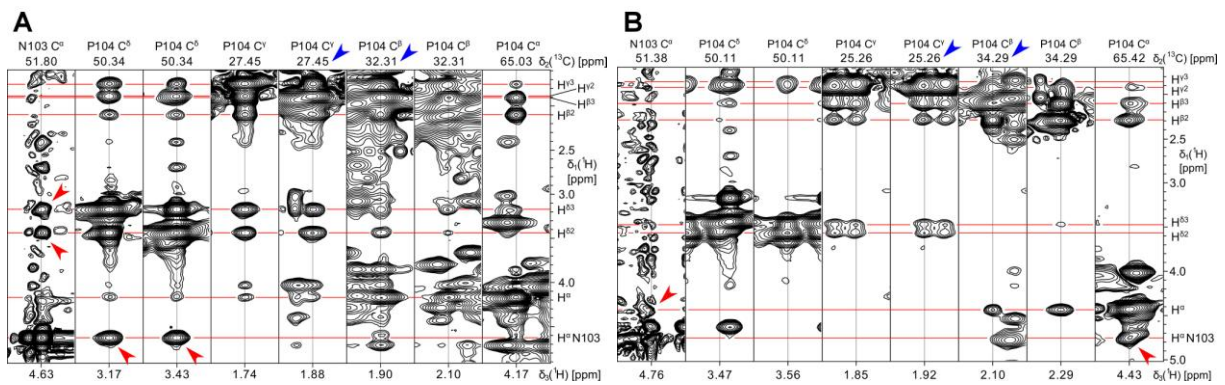
2. att. Proteīna NP_247299.1 KMR struktūra šķīdumā un salīdzinājums ar struktūru kristāliskā stāvoklī, stereoskati.

(A) Lentīšu attēlojums ar otrējās struktūras elementu identifikāciju.

(B) 20 KMR konformēru galvenās ķēdes atomu (brūni) superpozīcija ar struktūru kristāliskā stāvoklī (melna)



3. att. Kristalogrāfisko B-faktoru salīdzinājums ar atomu pozicionālo novirzi starp 20 KMR konformēriem. Augšējā panelī attēloti dati kristāliskai struktūrai un "atsauces" kristāliskai struktūrai, bet apakšējā panelī - KMR struktūrai un "atsauces" KMR struktūrai. Kristāliskai struktūrai B-faktori pārvērsti atomu pozicionālās novirzēs, izmantojot lineāru sakarību $\langle \Delta x \rangle = c \cdot \overline{B}$, kur koeficients c noteikts, veicot lineārās regresijas analīzi



4. att. Aminoskābju atlikuma Pro104 cis-trans izomerizācija. Sloksnes no 3D ^{13}C -izvērsta- $[\text{}^1\text{H}, \text{}^1\text{H}]$ -NOESY spektra, kas attēlo Asn103–Pro104 peptīdsaites trans- un cis-formas.

- (A) trans-Pro104 forma uzrāda $d_{\alpha\delta}^{NP}$ kodolu Overhauzera efektus (sarkanas bultiņas) un tipiskas $^{13}\text{C}^{\beta}$ un $^{13}\text{C}^{\gamma}$ ķīmiskās nobīdes (zilas bultiņas).
- (B) cis-Pro104 uzrāda $d_{\alpha\alpha}^{NP}$ kodolu Overhauzera efektus (sarkanas bultiņas) un stipri atšķirīgas $^{13}\text{C}^{\beta}$ un $^{13}\text{C}^{\gamma}$ ķīmiskās nobīdes (zilas bultiņas)

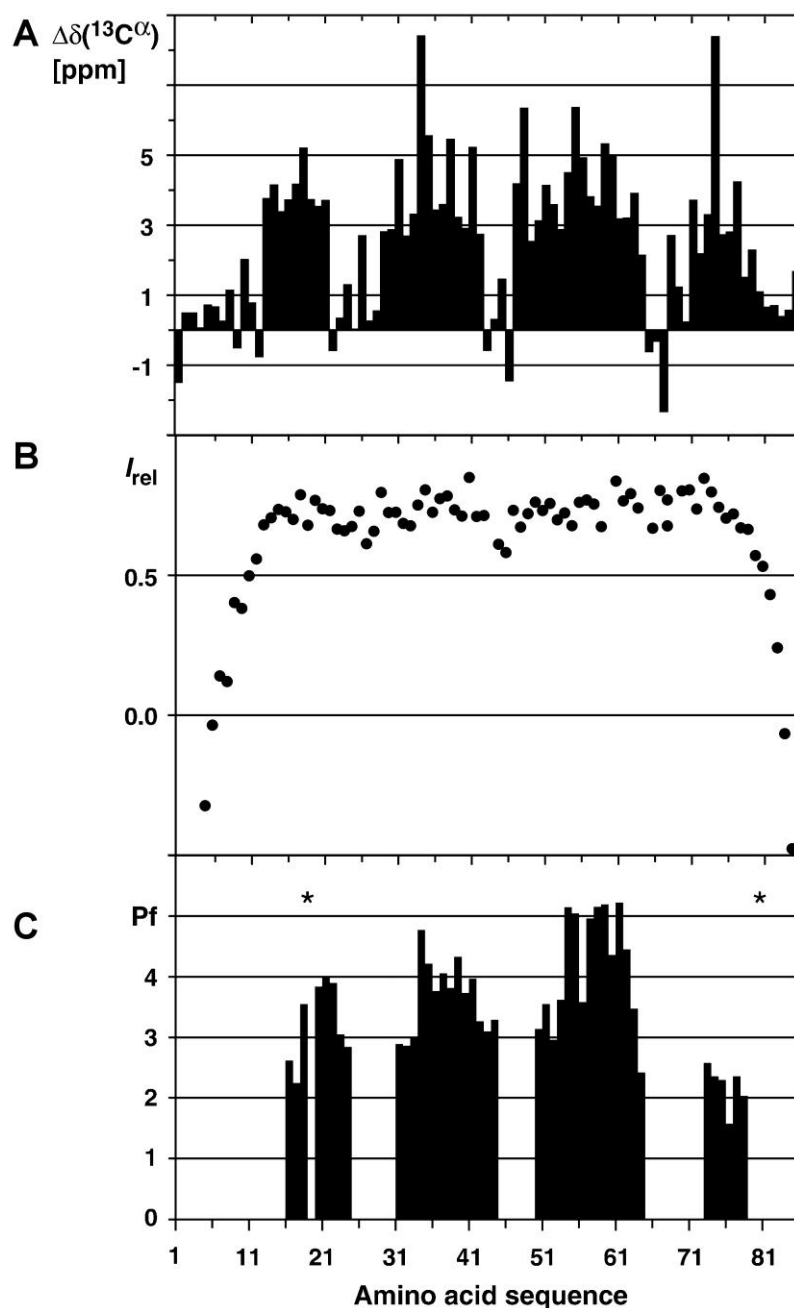
1.3. SARS KORONAVĪRUSA NESTRUKTURĀLĀ PROTEĪNA NSP7 KMR STRUKTŪRA ŠĶĪDUMĀ PIE pH 6.5

Smagā akūtā respiratorā sindroma koronavīrusa nestrukturālais proteīns 7 (Nsp7) izraisa interesi sakarā ar tā iespējamo lomu vīrusa RNS genoma transkripcijā un replikācijā. Uz tā bioloģisko nozīmīgumu norāda arī novērojums, ka Nsp7 gēna izslēgšana vai Nsp7/Nsp8 proteolītiskās šķelšanas vietas mutācija ir letāla peles hepatīta vīrusa (cits koronavīrusu paveids) gadījumā [14]. Nsp7 struktūra iepriekš noteikta ar KMR ūdens šķīdumā pie pH 7,5 [15], kā arī kristāliskā stāvoklī, izmantojot rentgenstaru difrakciju hetero-heksadekamēram kompleksam ar Nsp8 [16], kas pie tam uzrādījis RNS primāzes aktivitāti [17]. Kristāliskais komplekss satur astoņas molekulas Nsp7 un tikpat Nsp8, un veido kanālu ar RNS dubultspirālei atbilstošu platumu un lādiņu sadalījumu [16]. Abas iepriekš noteiktās Nsp7 struktūras satur četras spirāles kā vienīgus otrējās struktūras elementus, sastādot aptuveni 60% no 85 atlikumu garās polipeptīda ķēdes, tomēr to pozicionālais novietojums, garums un telpiskais izvietojums katrā struktūrā ir atšķirīgs, norādot uz iespējamu konformacionālo polimorfismu. Tādēļ nolēmām dziļāk izpētīt Nsp7 strukturālās variācijas un to iemeslus.

Projektu sākām, pārbaudot plašu klāstu šķīduma apstākļus, ar mērķi atrast vides nosacījumus, kas ļautu uzņemt augstas kvalitātes KMR spektrus un iegūt pilnīgāku konformacionālo ierobežojumu komplektu nekā bija iespējams pie pH 7,5 un augstas jonu koncentrācijas. Šķīdums ar pH 6,5 un zemu jonu koncentrāciju tika izvēlēts jaunai Nsp7 KMR struktūras noteikšanai. $^{13}\text{C}^{\alpha}$ ķīmiskās nobīdes liecināja, ka Nsp7 pie pH 6,5 satur četras α -spirāles (5. att. A) un vēlāk aprēķinātā struktūra uzrādīja četru antiparalēlu spirāļu kopumu (6. att.). Polipeptīda segments 1–10 noteiktu struktūru neveido, spirāles $\alpha 1$ (atlikumi 13–20) un $\alpha 2$ (atlikumi 29–42) savieno garāka, bet strukturāli labi definēta cilpa (atlikumi 21–28), savukārt spirāles $\alpha 2$ un $\alpha 3$ (atlikumi 47–65), kā arī $\alpha 3$ un $\alpha 4$ (atlikumi 70–82) savieno īsākas cilpas (atlikumi 43–46 un 66–69, respektīvi). Proteīna atsevišķu daļu dinamikas raksturošanai uzņēmām $^{15}\text{N}\{^1\text{H}\}$ -

NOE spektru, kas ir jutīgs pret polipeptīda galvenās ķēdes kustībām ps līdz ns laika skalā. Palielināts kustīgums novērojams N-gala dekaeptīda segmentam un C-gala heksapeptīdam, kas ietver spirāles $\alpha 4$ pēdējo vijumu (5. att. B), bet iegūtās NOE vērtības atlikumiem 13–79 liecina, ka šī centrālā polipeptīda daļa veido kompaktu globulāru struktūru. Lai gūtu ieskatu Nsp7 konformacionālajos līdzsvara procesos, noteicām proteīna galvenās ķēdes amīdgrupu $^1\text{H}/^2\text{H}$ apmaiņas ātrumus $^2\text{H}_2\text{O}$ šķīdumā (5. att. C, izteikti kā aizsarg-faktori (Pf) [18]) un veicām signālu formu analīzi. Amīdgrupu protonu aizsarg-faktori atspoguļo katra protona aizsardzības līmeni pret ķīmisko apmaiņu ar $^2\text{H}_2\text{O}$. Protoni, kas iesaistīti otrējās struktūras udeņraža saišu veidošanā vai savādāk nošķirti no šķīdinātāja uzrāda augstākus aizsarg-faktorus (jeb zemākus ķīmiskās apmaiņas ātrumus). Nsp7 proteīnā spirāles $\alpha 2$ un $\alpha 3$ uzrāda augstākos aizsarg-faktorus, tālāk seko $\alpha 1$ un $\alpha 4$, bet neregulāras struktūras polipeptīda segmentu aizsarg-faktori bija par zemiem, lai tos varētu izmērīt, izmantojot standarta metodi. Uzmanību piesaista spirāles $\alpha 4$ dati - salīdzinoši zemie aizsarg-faktori (5. att. C) un samazinātās $\Delta\delta(^{13}\text{C}^\alpha)$ vērtības atlikumiem 78–82 norāda uz pazeminātu $\alpha 4$ populāciju KMR paraugā, tai atrodoties dinamiskā līdzsvarā ar nestrukturētām, šķīdinātājam pieejamām konformācijām. Sub-nanosekundes laika skalas kustīgumu novērojām tikai C-gala heksapeptīda segmentam, kas liecina, ka $\alpha 4$ konformacionālo līdzsvaru visdrīzāk nosaka lēnākas kustības. To apliecināja signālu formu analīze - atlikumi Ser 63 un Gln 65 $\alpha 3$ C-gala tuvumā, Val 68 $\alpha 3$ un $\alpha 4$ savienojošajā cilpā, kā arī Asn 71 un Leu 73 $\alpha 4$ sākumā visi uzrāda izteiktu signālu paplašināšanos, salīdzinot ar atlikumiem, kas nav iesaistīti $\alpha 4$ konformacionālajā līdzsvarā.

Salīdzinot šajā darbā noteikto Nsp7 struktūru ar iepriekš noteiktajām, varam secināt, ka struktūras kodolu veidojošo spirāļu $\alpha 2$ un $\alpha 3$ garumi, pozicionālais un telpiskais novietojums saglabājas, bet spirāles $\alpha 1$ un $\alpha 4$ dažādās vidēs pieņem atšķirīgu garumu, novietojumu un relatīvo orientāciju (7. att.). Nsp7-Nsp8 kompleksa kristāliskā struktūra no pH 6,5 KMR struktūras atšķiras galvenokārt ar $\alpha 4$ spirāles rotāciju prom no $\alpha 2/\alpha 3$ kodola un polipeptīda segmenta 3–12 salocīšanos, veidojot garāku $\alpha 1$ (7. att. A), kas gan neatbilst šķīdumā noteiktajiem amīdgrupu protonu aizsarg-faktoriem (5. att. C). Garākā kristāliskās struktūras $\alpha 1$ tad ieņem pH 6,5 KMR struktūras $\alpha 4$ vietu (7. att. A). Savukārt pH 7,5 KMR struktūrā $\alpha 4$ piekļaujas tikai $\alpha 3$ un ar $\alpha 2$ kontaktus neveido, tādējādi spirāles $\alpha 2$, $\alpha 3$ un $\alpha 4$ novietojas vienā plaknē un $\alpha 1$ piekļaujas $\alpha 2$ un $\alpha 3$ aptuveni 45° leņķī (7. att. B). Tas liek domāt, ka Nsp8 saistās ar Nsp7 īslaicīgi eksistējošu formu (līdzsvara stāvokli), kurā spirāle $\alpha 4$ nav strukturēta (5. att. A un C); Nsp8 saistīšanās izraisa īsākas spirāles $\alpha 4$ veidošanos, kā arī garākas spirāles $\alpha 1$ veidošanos, kas ieņem pH 6,5 KMR struktūras $\alpha 4$ vietu (7. att. A). Kristāliskā struktūrā $\alpha 1$ un $\alpha 4$ novietojumu stabilizē vairāki hidrofobi kontakti ar Nsp8 sānu ķēdēm. Interesanti, ka spirāļu $\alpha 1$ līdz $\alpha 3$ aminoskābju atlikumi ir labi saglabājušies visos koronavīrusos, bet $\alpha 4$ atlikumi ir nepastāvīgi un tādējādi varētu noteikt individuālu sugu proteīnu specifiskumu un fizioloģisko aktivitāti. Šie secinājumi varētu tikt izmantoti jaunu pretvīrusu zāļu vielu radīšanā - Nsp7 spirāles $\alpha 4$ ieslēgšana konformācijā, kas nevar piesaistīt Nsp8 varētu kaitēt koronavīrusu replikācijas mehānismam.

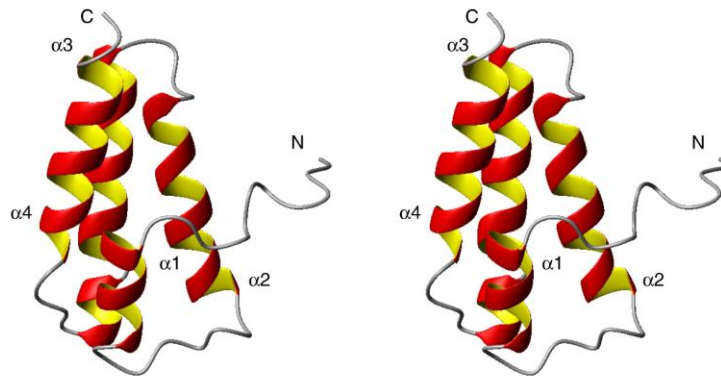


5. att. KMR datu grafiki pret aminoskābju secību.

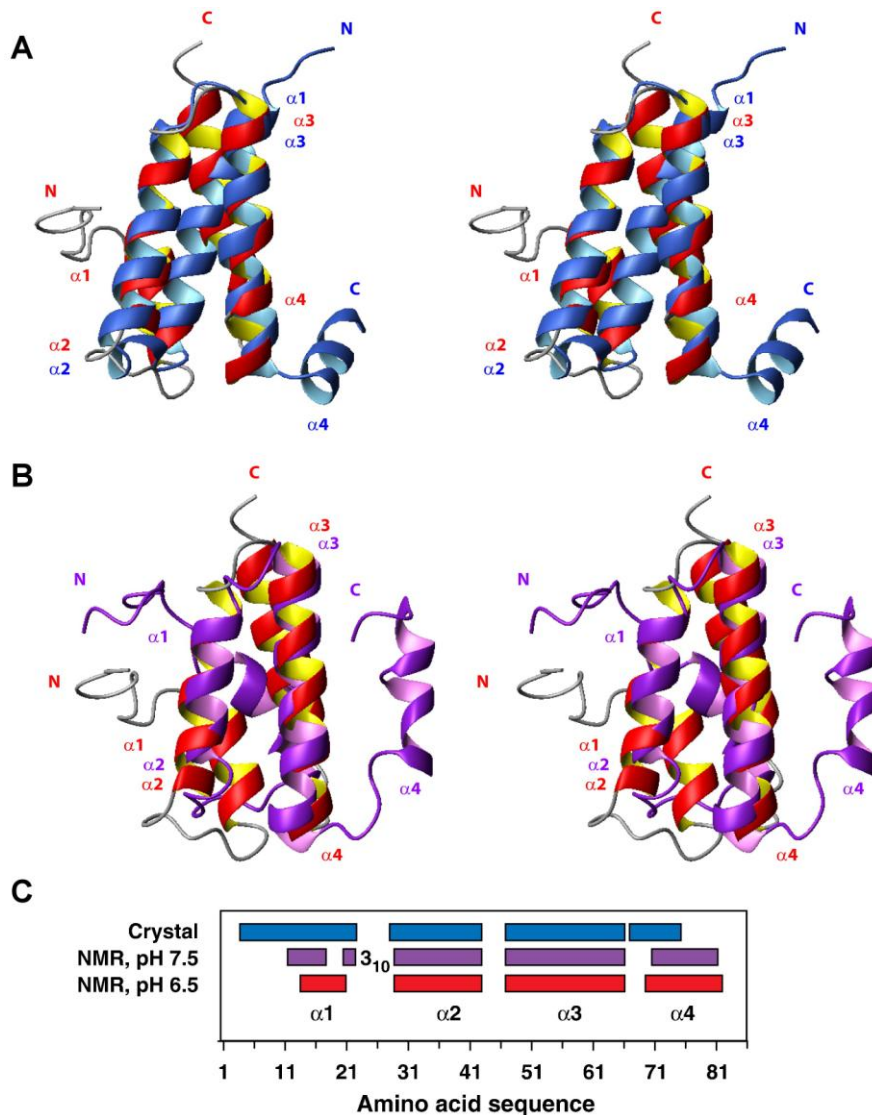
(A) $^{13}\text{C}^{\alpha}$ ķīmisko nobīžu novirzes no nejaušas ķēdes ("random coil") vērtībām, $\Delta\delta(^{13}\text{C}^{\alpha})$.

(B) $^{15}\text{N}\{^1\text{H}\}$ -NOE vērtības, I_{rel} .

(C) Amīdgrupu $^1\text{H}/^2\text{H}$ apmaiņas aizsarg-faktori, Pf. Zvaigznītes norāda atlikumus, kuru Pf vērtības nevarēja precīzi noteikt sakarā ar signālu pārklāšanos [^{15}N , ^1H]-HSQC spektrā



6. att. Nsp7 stuktūra šķīdumā pie pH 6,5. Lentīšu attēlojums ar otrējās struktūras elementu identifikāciju, stereoskats



7. att. Nsp7 struktūras šķīdumā pie pH 6,5 (sarkani-dzeltena) superpozīcija ar iepriekš noteiktajām Nsp7 struktūrām, stereoskatī. Struktūru superpozīcija veikta spirāļu $\alpha 2$ un $\alpha 3$ aminoskābju atlikumu galvenajām ķēdēm.

(A) Superpozīcija ar Nsp7-Nsp8 kompleksa kristālisko struktūru (zila).

(B) Superpozīcija ar Nsp7 struktūru šķīdumā pie pH 7,5 (violeta).

(C) Otrējo struktūras elementu atrašanās vietas visās eksperimentāli noteiktajās Nsp7 struktūrās

2. NODAĻA. PROTEĪNU-LIGANDU MIJIEDARBĪBU PĒTĪJUMI AR KMR

KMR spektroskopija ir ļoti piemērota fizikālā metode mazu molekulu un proteīnu mijiedarbības noteikšanai un kvantificēšanai. Tā var tikt izmantota plašu potenciālo zāļu vielu bibliotēku skrīningam, kompleksu disociācijas konstanšu mērījumiem, mijiedarbības vietas uz proteīna virsmas noteikšanai vai pat proteīna-liganda kompleksa trīsdimensionālās struktūras noskaidrošanai.

Proteīnu-ligandu mijiedarbību pētījumiem izmantojamās KMR metodes var iedalīt proteīna signālus novērojošās un liganda signālus novērojošās. Proteīna signālus novērojošās metodes dod iespēju precīzi noteikt liganda mijiedarbības vietu uz proteīna virsmas, bet tām nepieciešams liels daudzums proteīna (vēlams ^{15}N iezīmēta). Šo metožu pamatā ir liganda inducētas proteīna atomu lokālās magnētiskās vides izmaiņas, ko novēro kā signālu nobīdi spektrā. To izmantošana aprakstīta R3H domēna un mononukleotīdu kompleksu pētījumos. Liganda signālus novērojošās piesātinājuma pārnese starpības metodes (Saturation Transfer Difference, STD) izmantošana aprakstīta CrAT pētījumos. Šīs metodes pamatā ir tieša starpmolekulāra magnetizācijas pārnese starp proteīna un liganda molekulām. Ligandu novērojošām metodēm nepieciešams mazāks, pie tam neiezīmēta proteīna daudzums, kas ļauj veikt daudz plašāku ligandu skrīningu, bet iegūtā informācija par saistības vietu ir trūcīgāka.

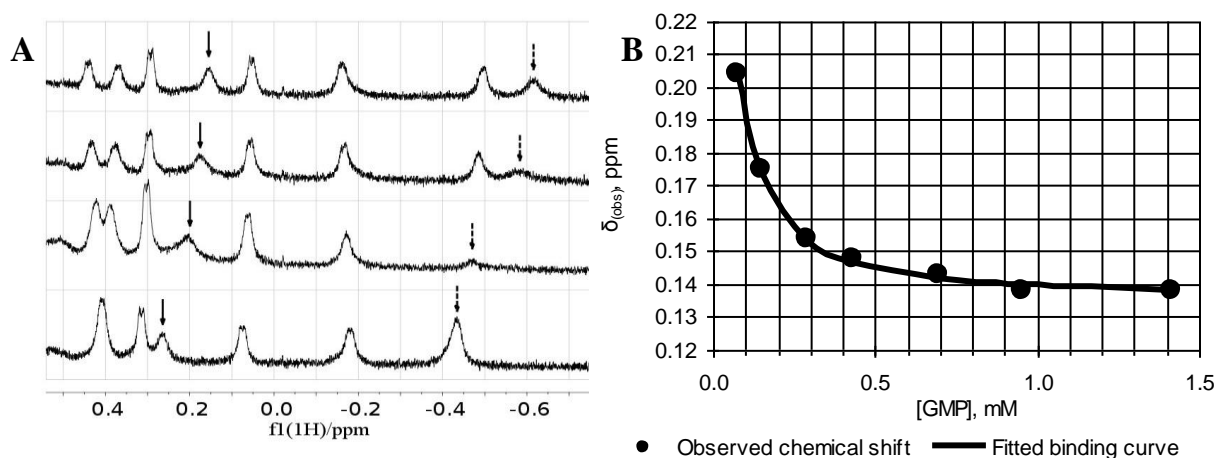
2.1. DOMĒNA R3H KOMPLEKSI AR MONONUKLEOTĪDIEM

R3H domēns iepriekš atklāts kā vienojošs sekvenču motīvs šķietami nukleīnskābes-saistošos proteīnos no dažādiem organismiem [19]. Tas nosaukts pēc nemainīga arginīna un starp sugām labi saglabāta histidīna atlikumu klātbūtnes, kas aminoskābju sekvencē atdalīti ar trim atlikumiem. R3H domēna no cilvēka Sūbp-2 proteīna trīsdimensionālā struktūra šķīdumā iepriekš noteikta ar KMR spektroskopiju [20]. Šajā darbā atspoguļoti rezultāti no R3H domēna saistības pētījumiem ar mononukleotīdiem, izmantojot KMR, kas veikti ar mērķi noteikt R3H domēna saistības partneru izvēli.

Literatūrā ir vairākas KMR metodes, kas pielietojamas mazu molekulu un proteīnu saistības konstatēšanai, kā arī kvantificēšanai [21]. Mēs izvēlējāmies novērot proteīna metilgrupu ^1H KMR ķīmisko nobīžu izmaiņas attiecībā pret mazo molekulu koncentrāciju. Molekulai, kas attiecībā pret KMR laika skalu ātri apmainās starp brīvo un saistīto formu, jebkurš KMR parametrs ir moldaļas svērtais vidējais no vērtībām, ko uzrāda katra forma individuāli. Tādēļ, uzņemot spektrus pie dažādām proteīna un mazo molekulu koncentrāciju attiecībām un veicot simulētu datu nelineāru mazāko kvadrātu aproksimāciju pret eksperimentāliem datiem, iespējams noteikt kompleksa disociācijas konstanti (K_D). R3H domēna KMR titrēšanas ar guanozīna monofosfātu (GMP) rezultāti redzami 8. att. (A). Disociācijas konstantes noteikšanas ilustrācijai izmantots signāls pie 0.26 ppm. 8. att. (B) attēlota novērotās ķīmiskās nobīdes atkarība no GMP koncentrācijas, kā arī simulētā saistīšanās izoterma, kas nosaka K_D . Mononukleotīdu, kuriem konstatējam saistību ar R3H domēnu, K_D vērtības apkopotas 1. tabulā.

Iespējamās DNS saistības vietas identifikācijai uz R3H virsmas, novērojām tā galvenās ķēdes amīdgrupu ķīmisko nobīžu izmaiņas 2D ^{15}N -HSQC spektrā, paraugam pievienojot dGMP. Lielākā daļa mainīgo signālu piederēja aminoskābju atlikumiem no spirāles $\alpha 2$ un virknēm $\beta 2$ un $\beta 3$, kas atrodas domēna vienā pusē un ietver R3H

motīvu, apstiprinot tā funkcionālo lomu DNS saistīšanās. Iegūtie rezultāti ļāva izveidot modeli R3H domēna mijiedarbībām ar dGMP, kas vēlāk tika apstiprināts, nosakot R3H kompleksa ar dGMP trīsdimensionālo struktūru (9. att.).



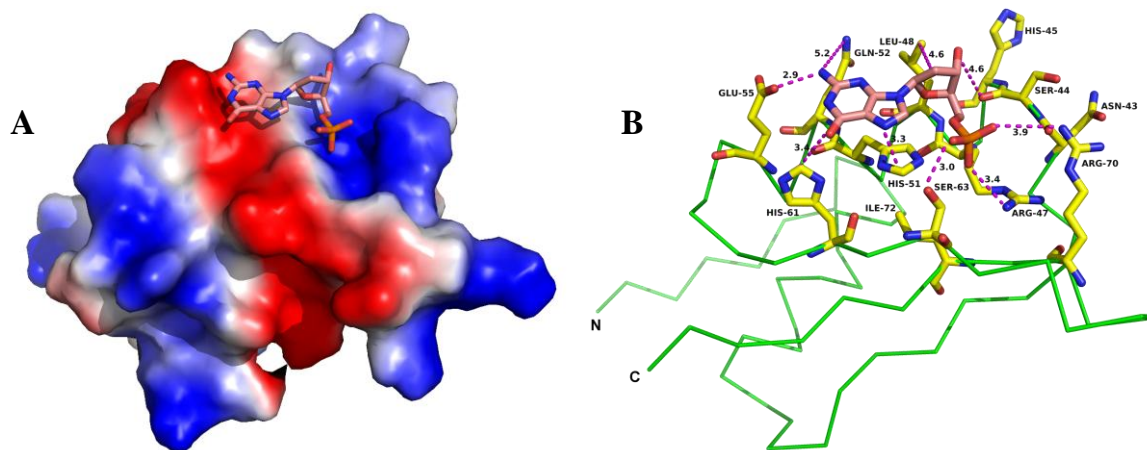
8. att. (A) R3H domēna ¹H KMR spektra metilgrupu rajons pie dažādām R3H un GMP koncentrāciju attiecībām. Bultiņas norāda uz signāliem, kas izmantoti K_D noteikšanai. (B) R3H metilgrupas signāla pie 0.26 ppm atkarība no GMP koncentrācijas (mM) pie konstantas R3H domēna koncentrācijas (0.1 mM). Apļi attēlo eksperimentālos punktus un līkne ir simulētā saistīšanās izoterma

1. tabula

R3H domēna kompleksu ar mononukleotīdiem disociācijas konstantes

N ^o	Nukleotīds (saīsinājums)	K _D , μM	N ^o	Nukleotīds (saīsinājums)	K _D , μM
1	GMP	33.6 ± 1.0	5	dGMP	8.0 ± 0.1
2	AMP	109.0 ± 4.0	6	dAMP	58.7 ± 4.7
3	UMP	106.4 ± 10.6	7	dTMP	268.5 ± 54.5
4	CMP	290.5 ± 30.5	8	dCMP	276.0 ± 43.0

Iegūtie rezultāti parāda, ka R3H domēns saista mononukleotīdu monofosfātus, veidojot kompleksus ar disociācijas konstanti mikro-molārā rajonā. Saistība ar dezoksiribo-mononukleotīdiem ir ievērojami stiprāka nekā ar ribo-mononukleotīdiem, īpaši to nukleotīdu gadījumā, kas satur purīna atlikumu. R3H domēns nesaistās ar mononukleotīdu difosfātiem un trifosfātiem, kā arī nefosforilētiem nukleozīdiem, kas tāpat nav sastopami DNS un RNS. R3H domēns nesaista arī 3'-fosforilētus mononukleotīdus. Dezoksi-guanozīna monofosfāta saistības vieta uz R3H identificēta pēc R3H signālu nobīdēm 2D ¹⁵N-HSQC spektrā. Saistīšanās notiek elektrostātisku mijiedarbību ar Arg47 un Arg70 sānu ķēdēm, π-π mijiedarbībām ar His51 un vairākām starpmolekulārām ūdeņraža saitēm (9. att. B). His51 atlikums varētu būt atbildīgs par selektivitāti pret purīna-saturošiem nukleotīdiem, purīnu gadījumā veidojot labākas π-sistēmu mijiedarbības, bet Leu48 atlikuma hidrofobiskās sānu ķēdes tuvums nukleotīda ribozes atlikumam varētu nodrošināt dezoksiribonukleotīdu selektivitāti. Iegūtie rezultāti norāda, ka R3H domēns mijiedarbojas ar DNS, un varētu būt selektīvs pret guanozīna bagātām DNS sekvencēm.



9. att. R3H domēna kompleksa ar dGMP trīsdimensionālā struktūra šķīdumā. (A) dGMP novietojums uz R3H molekulārās virsmas. Virsma iekrāsota, attēlojot tās elektrostatisko potenciālu. Negatīvs potenciāls iekrāsots sarkanā krāsā, bet pozitīvs - zilā.

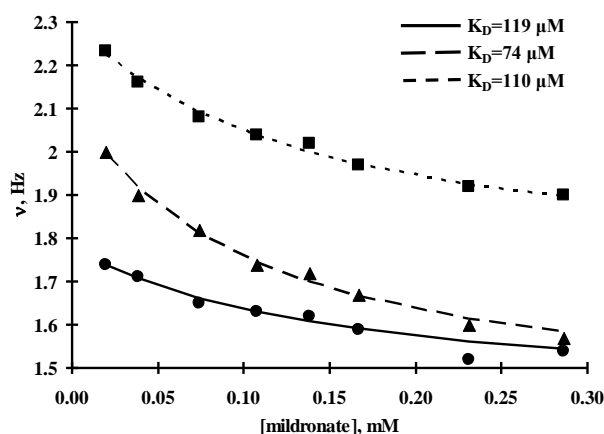
(B) dGMP (iekrāsots rozā) saistības vietas pietuvinājums. Aminoskābju atlikumi, kas atrodas 5Å attālumā no liganda ir identificēti un attēloti ar dzeltenām nūjiņām.

Slāpekļa un skābekļa atomi iekāšoti attiecīgi zilā un sarkanā krāsā. Ūdeņraža atomi nav attēloti. Potenciālās mijiedarbības uzziņētas ar pārtrauktām līnijām, tām norādīti attālumi starp atbildīgajām grupām

2.2. KARNITĪNA ACETILTRANSFERĀZES INHIBĪCIJA AR ENERĢIJAS METABOLISMA REGULATORU MILDRONĀTU

Karnitīna acetiltransferāze (CrAT) katalizē atgriezenisko acetilgrupas pārnese reakciju starp acetil-koenzīmu A (acetil-CoA) un L-karnitīnu [22], tādējādi regulējot brīva CoA un aktivētu acetilgrupu koncentrāciju šūnās [23]. Mildronāts (3-(2,2,2-trimetilhidrazīnija)-propionāts) ir kardioprotektīvs līdzeklis, kura darbības pamatā ir karnitīna koncentrācijas samazināšana un brīvo taukskābju β -oksidēšanas inhibēšana [24-25]. Mildronāts ir strukturāli līdzīgs karnitīnam un īpaši tā biopreksoram γ -butirobetaīnam. Šajā darbā, izmantojot bioķīmiskās metodes, KMR spektroskopiju un molekulāro modelēšanu pētīta kardioprotektīvā līdzekļa mildronāta un CrAT mijiedarbība.

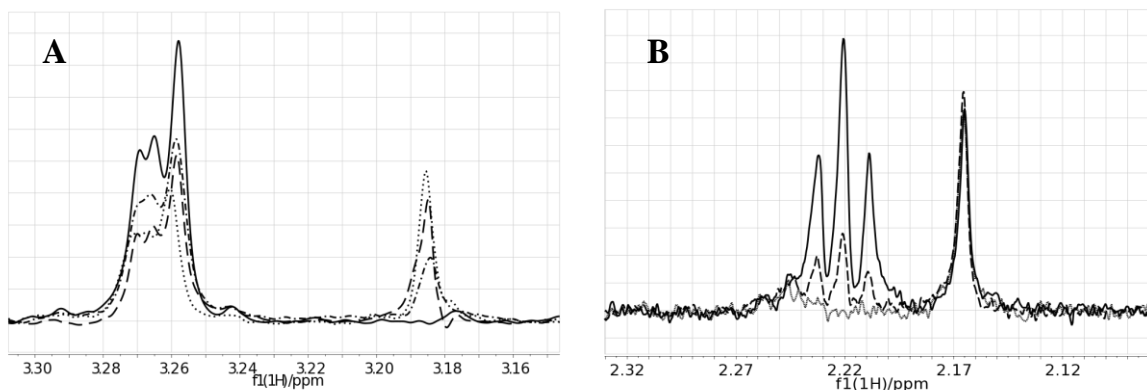
Mildronāta inhibitorās efektivitātes raksturošanai noteicām mildronāta un CrAT kompleksa disociācijas konstanti, novērojot mildronāta ^1H KMR signālu pusplatumu izmaiņas atkarībā no saistības partneru koncentrāciju attiecības. Sakarā ar ļoti nelielām novērotajām pusplatumu izmaiņām (aptuveni 0,5 Hz) un matemātiskās metodes augsto jutību pret mērījumu kļūdu, noteiktos pusplatumus normalizējām attiecībā pret iekšējo standartu (DSS), tādējādi kompensējot magnētiskā lauka nehomogenitātes ietekmi. Kompleksa disociācijas konstantes aprēķināšanai veicām saistīšanās izotermu simulāciju, pielāgojot divus parametrus - kompleksa disociācijas konstanti K_D un saistītās formas pusplatumu $v_{\text{saist.}}$. Mildronāta ^1H KMR signālu pusplatumu saistīšanās ar CrAT izotermas attēlotas 10. attēlā. Aprēķinātā kompleksa disociācijas konstante $K_D = 101 \pm 19 \mu\text{M}$.



10. att. Mildronāta trimetilamonija grupas (apļi), α -metilēngrupas (trijstūri) un β -metilēngrupas (kvadrāti) ^1H KMR signālu pusplatumi atkarībā no mildronāta koncentrācijas (mM) pie nemainīgas CrAT koncentrācijas (20 μM). Apļi, kvadrāti un trijstūri attēlo eksperimentālos punktus un līknes ir simulētās saistīšanās izoterms

Mildronāta saistīšanās vietas identifikācijai izmantojām konkurējošo ligandu piesātinājuma pārneses diferences KMR metodi [26], kuras pamatā ir indikator-liganda signālu samazināšanās vai izžušana pēc konkurējoša liganda pievienošanas proteīna un indikator-liganda paraugam. Konkurējošais ligands aizvieto indikator-ligandu tā saistības vietā, tāpēc uz indikator-ligandu saistītajā stāvoklī no proteīna tiek pārnesti mazāks daudzums magnetizācijas, ko vēlāk detektē. Indikator-liganda signālu samazināšanās pakāpe atkarīga no konkurējošā liganda afinitātes pret proteīnu, tāpēc eksperiments dod arī aptuvenu informāciju par konkurējošā liganda saistīšanās tieksmi relatīvi pret indikator-ligandu. Pakāpeniska mildronāta pievienošana karnitīna-CrAT kompleksam izraisa ievērojamu un nepārtrauktu karnitīna signālu samazināšanos un mildronāta signālu pieaugumu piesātinājuma pārneses diferences KMR spektrā (11. att. A), norādot, ka mildronāts saistās karnitīna saistības vietā, to izspiežot. Šie dati arī apstiprināja noteiktās mildronāta-CrAT kompleksa disociācijas konstantes vērtību, parādot, ka mildronāta un karnitīna konstantes ir aptuveni līdzvērtīgas un atrodas mM rajonā. Izmantojot analoģu eksperimentu parādījām, ka mildronāts nekonkurē ar acetil-CoA un neaizņem tā saistīšanās vietu (11. att. B).

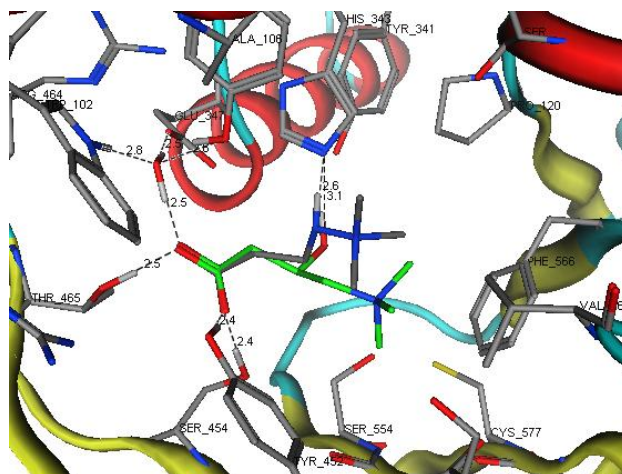
Piesātinājuma pārneses diferences KMR spektri var arī tikt izmantoti aptuvenai saistītās liganda konformācijas noteikšanai. Karnitīna α -CH₂, γ -CH₂ un trimetilamonija grupu signālu piesātinājuma pārneses palielinājuma faktori (respektīvi 17%, 8.5% un 37%) labi atbilst karnitīna konformācijai karnitīna un CrAT kompleksa rentģestaru difrakcijas kristāliskā struktūrā: γ -CH₂ grupa vērsta pret šķīdinātāju, rezultātā uzrādot vismazāko piesātinājuma pārnesi no proteīna, kamēr trimetilamonija grupa novietojas hidroforā kabatā un atrodas ciešā kontaktā ar proteīna virsmu, iegūstot vislielāko piesātinājuma pārnesi no proteīna. Mildronāta α -CH₂, β -CH₂ un trimetilamonija grupu signālu piesātinājuma pārneses palielinājuma faktori (respektīvi 11%, 17% un 1%) norāda, ka mildronātam saistītā stāvoklī ir atšķirīgs trimetilamonija grupas novietojums, salīdzinot ar karnitīnu un ļoti iespējams, ka šī grupa vērsta pret šķīdinātāju. Molekulārās modelēšanas rezultāti apstiprināja šādu mildronāta saistītā stāvokļa konformāciju, kur trimetilamonija grupa vērsta pret šķīdinātāju (12. att.).



11. att. (A) karnitīna γ -metilēngrupas piesātinājuma pārnesešes diferences KMR spektra signāls (pie 3.265 ppm) samazinās un mildronāta trimetilamonija grupas signāls (pie 3.185 ppm) pieaug, karnitīna un CrAT paraugam pievienojot konkurējošo ligandu mildronātu.

(B) acetil-CoA signāla (pie 2.165 ppm) intensitāte nemainās, acetil-CoA un CrAT paraugam pievienojot nekonkurējošo ligandu mildronātu. Mildronāta α -metilēngrupas signāls (pie 2.22 ppm) pieaug, palielinoties mildronāta koncentrācijai

Iegūtie rezultāti parāda, ka mildronāts CrAT inhibē konkurējoši, bloķējot karnitīna saistīšanās vietu, bet ne acetil-CoA saistīšanās vietu. Mildronāta konformācija saistītā stāvoklī līdzinās karnitīnam, izņemot trimetilamonija grupas orientāciju, kas mildronāta molekulā vērsta uz šķīdinātāju. Mildronāta un CrAT kompleksa disociācijas konstante ir apmēram 0,1 mM, bet inhibīcijas konstante 1,6 mM. Rezultāti liecina, ka mildronāta kardioprotektīvā iedarbība varētu būt daļēji saistīta ar CrAT inhibīciju un sekojošu šūnu enerģijas metabolisma ceļu regulāciju.



12. att. Karnitīna novietojums CrAT kompleksa kristāliskā struktūrā un salīdzinājums ar modelēto mildronāta novietojumu. Karnitīna oglekļa atomi iekrāsoti zaļā krāsā, bet mildronāta - baltā, ūdeņraža atomi nav attēloti.

SECINĀJUMI

Izmantojot modernas KMR tehnikas, pirmo reizi noteiktas proteīnu DnaI-N, NP_247299.1 un Nsp7 pie pH 6,5 trīsdimensionālās struktūras šķīdumā, kā arī atkārtota struktūras noteikšana R3H domēnam kompleksā ar dGMP. Struktūru noteikšanai izmantoti proteīnu paraugi ar ^{15}N vai dubultu ^{15}N un ^{13}C izotopu iezīmi, kas iegūti rekombinantās ekspresijas ceļā. Iegūtās ķīmiskās nobīdes un struktūru koordinātas ievietotas attiecīgi Bioloģiskās magnētiskās rezonanses datu bankā un Proteīnu datu bankā, kas ir publiski pieejamas.

DnaI:

Pirmoreiz noteiktā DnaI-N proteīna struktūra satur četras α -spirāles, divas β -virknes, kā arī īsu 3_{10} spirāli un uzrāda jaunu telpiskās uzbūves tipu. DnaI satur cinka jonu, kam visdrīzāk ir strukturāli nozīmīga loma. Konservatīvu, hidrofobisku aminoskābju atlikumu klātbūtne struktūru neveidojošajā N-gala fragmentā liecina par to funkcionālu lomu, iespējams mijiedarbībā ar helikāzi.

NP_247299.1:

Demonstrēta automatizētās proteīnu struktūras noteikšanas pielietojamība, kas ļauj ievērojami samazināt instrumenta noslodzes un spektroskopista laiku, kas nepieciešams spektru uzņemšanai, analīzei un datu precizēšanai. Izstrādāts jauns protokols proteīnu rentgenstaru difrakcijas un KMR struktūru analīzei un salīdzināšanai, kas ļauj interpretēt lokāla rakstura atšķirības globāli ļoti līdzīgajās proteīna NP_247299.1 struktūrās. Lokāla rakstura atšķirības novēro segmentiem ar augstiem kristalogrāfiskajiem B faktoriem, kas korelē ar dinamisku struktūrālu polimorfismu šķīdumā.

Nsp7:

Nsp7 proteīns veido 4 α -spirāles, no tām struktūras kodolu veidojošo spirāļu $\alpha 2$ un $\alpha 3$ garumi, pozicionālais un telpiskais novietojums dažādās vidēs saglabājas, bet spirāles $\alpha 1$ un $\alpha 4$ pieņem atšķirīgu garumu, novietojumu un relatīvo orientāciju. Atlikumi 70–82 atrodas konformacionālā līdzsvarā starp α -spirāles ($\alpha 4$) un nestrukturētu formu. Tas liek domāt, ka Nsp8 proteīns saistās ar Nsp7 īslaicīgi eksistējošu formu (līdzsvara stāvokli), kurā spirāle $\alpha 4$ nav strukturēta; Nsp8 saistīšanās izraisa īsākas spirāles $\alpha 4$ veidošanos, kā arī garākas spirāles $\alpha 1$ veidošanos, kas ieņem pH 6,5 KMR struktūras $\alpha 4$ vietu. Kristāliskā struktūrā $\alpha 1$ un $\alpha 4$ novietojumu stabilizē vairāki hidrofobi kontakti ar Nsp8 sānu ķēdēm.

R3H:

Atrasts, ka R3H domēns saista mononukleotīdu monofosfātus, veidojot kompleksus ar disociācijas konstanti mikro-molārā rajonā. Saistība ar dezoksiribonukleotīdiem ir ievērojami stiprāka nekā ar ribonukleotīdiem, īpaši to nukleotīdu gadījumā, kas satur purīna atlikumu. R3H domēns nesaistās ar mononukleotīdu difosfātiem un trifosfātiem, kā arī nefosforilētiem nukleozīdiem, kas tāpat nav sastopami DNS un RNS. R3H domēns nesaista arī 3'-fosforilētus mononukleotīdus. Dezoksi-guanozīna monofosfāta saistīšanās pie R3H notiek pateicoties elektrostatiskām mijiedarbībām ar Arg47 un Arg70 sānu ķēdēm, π - π mijiedarbībām ar His51 un vairākām starpmolekulārām ūdeņraža saitēm. His51

atlikums varētu būt atbildīgs par selektivitāti pret purīna-saturošiem nukleotīdiem, purīnu gadījumā veidojot labākas π -sistēmu mijiedarbības, bet Leu48 atlikuma hidrofobiskās sānu ķēdes tuvums nukleotīda ribozes atlikumam varētu nodrošināt dezoksiribonukleotīdu selektivitāti.

CrAT:

Izveidots modelis antiišēmiskā līdzekļa mildronāta un CrAT mijiedarbībai. Mildronāts CrAT inhibē konkurējoši, bloķējot karnitīna saistīšanās vietu, bet ne acetil-CoA saistīšanās vietu. Mildronāta konformācija saistītā stāvoklī līdzinās karnitīnam, izņemot trimetilamonija grupas orientāciju, kas mildronāta molekulā vērsta uz šķīdinātāju. Mildronāta un CrAT kompleksa disociācijas konstante ir apmēram 0,1 mM, bet inhibīcijas konstante 1,6 mM. Rezultāti liecina, ka mildronāta kardioprotektīvā iedarbība varētu būt daļēji saistīta ar CrAT inhibīciju un sekojošu šūnu enerģijas metabolisma ceļu regulāciju.

PILNS AUTORA PUBLIKĀCIJU SARAKSTS

1. Johnson M.A., Jaudzems K., Wüthrich K. NMR structure of the SARS coronavirus nonstructural protein Nsp7 in solution at pH 6.5// *Journal of Molecular Biology*. - 2010. - Vol. 402. - p. 619-628.
2. Jaudzems K., Geralt M., Serrano P., Mohanty B., Horst R., Pedrini B., Elsliger M.-A., Wilson I.A., Wüthrich K. NMR structure of the protein NP_247299.1: comparison with the crystal structure// *Acta Crystallographica Section F*. - 2010. - Vol. 66. - p. 1367-1380.
3. Mohanty B., Serrano P., Pedrini B., Jaudzems K., Geralt M., Horst R., Herrmann T., Elsliger M.-A., Wilson I.A., Wüthrich K. Comparison of NMR and crystal structures for the proteins TM1112 and TM1367// *Acta Crystallographica Section F*. - 2010. - Vol. 66. - p. 1381-1392.
4. Serrano P., Pedrini B., Geralt M., Jaudzems K., Mohanty B., Horst R., Herrmann T., Elsliger M.-A., Wilson I.A., Wüthrich K. Comparison of NMR and crystal structures highlights conformational isomerism in protein active sites// *Acta Crystallographica Section F*. - 2010. - Vol. 66. - p. 1393-1405.
5. Jaudzems K., Kuka J., Gutsaits A., Zinovjevs K., Kalvinsh I., Liepinsh E., Liepinsh E., Dambrova M. Inhibition of carnitine acetyltransferase by mildronate, a regulator of energy metabolism// *Journal of Enzyme Inhibition and Medicinal Chemistry*. - 2009. - Vol. 24. - p. 1269-1275.
6. Jaudzems K., Zhulyenkov D., Liepinsh E. R3H domain complexes with mononucleotides// *Scientific Journal of RTU. Material Science and Applied Chemistry*. - 2009. - Vol. 20. - p. 58-63.
7. Loscha K.V., Jaudzems K., Ioannou C., Su X.-C., Hill F.R., Otting G., Dixon N.E., Liepinsh E. A novel zinc-binding fold in the helicase interaction domain of the *Bacillus subtilis* DnaI helicase loader// *Nucleic Acids Research*. - 2009. - Vol. 37. - p. 2395-2404.
8. Zinovjevs K., Jaudzems K., Liepinsh E. Determination of the 3D structure of DNA hexamers (ATGCAT, ATCGAT, TAGCTA, TACGTA) by 2D NMR spectroscopy// *Scientific Journal of RTU. Material Science and Applied Chemistry*. - 2008. - Vol. 16. - p. 99-106.
9. Jaudzems K., Liepinsh E., Bizdena E. Determination of the configuration of chiral centres in the derivatives of monosaccharides by NMR// *Scientific Journal of RTU. Material Science and Applied Chemistry*. - 2006. - Vol. 13. - p. 71-77.

PATEICĪBA

Šī darba izstrādāšana nebūtu iespējama bez daudzu citu cilvēku līdzdalības. Es vēlos pateikties šiem cilvēkiem par viņu nesavtīgo darbu, sadarbību un dalīšanos ar zināšanām.

Pirmkārt, es vēlos pateikties šī darba zinātniskajam vadītājam prof. Edvardam Liepiņam par manis izskološanu no zinātkāra maģistrantūras studenta līdz patstāvīgam zinātniekam ar pasaules līmeņa pieredzi. Tikai pateicoties viņa mērķtiecīgiem centieniem un ietekmei man radās iespēja divus gadus stažēties Nobela prēmijas laureāta, prof. Kurt Wüthrich laboratorijā ASV. Paldies arī par viņa palāvību, uzticot man arvien atbildīgākus uzdevumus, kas lieliski stimulēja manu interesi un entuziasmu.

Otrkārt, es vēlos pateikties manam zinātniskajam vadītājam Skripsa pētniecības institūtā ASV, prof. Kurt Wüthrich par pirmklasīgu darba vadīšanu, iedvesmojošām diskusijām un uzmanību, ko saņēmu. Es augsti vērtēju man doto iespēju strādāt tik izcilas personības vadībā.

Īpašs paldies maniem kolēģiem Skripsa pētniecības institūtā Michael Geralt, Pedro Serrano, Reto Horst, Biswaranjan Mohanty, Pawel Stanczak, Bill Pedrini, Maggie Johnson, Lukas Susac, Atia-tul Wahab, Amarnath Chatterjee un Sharron Thomas par atbalstu un aizrautīgām diskusijām, kas radīja lielisku atmosfēru auglīgam darbam. Viņi man kļuva par tuviem draugiem.

Tikpat īpašs paldies maniem kolēģiem Latvijas Organiskās sintēzes institūtā (LOSI) Dr. Osvaldam Pugovičam, Dacei Kārklei, Dr. Aigaram Jirgenonam un daudziem citiem par sadarbību, atbalstu manām studijām ASV un mani saistošu pētījumu virzienu attīstībai LOSI. Esot ASV, es vienmēr jutu, ka tieku gaidīts atpakaļ. Paldies viņiem par to.

Es gribētu izteikt pateicību visiem cilvēkiem, ar kuriem sadarbojos darbā iekļauto publikāciju izstrādē – prof. Gottfried Otting grupai, prof. Nick Dixon grupai, prof. Kurt Wüthrich grupai, Dr. Maijas Dambrovas grupai un Dmitrijam Žuļenkovam.

Es gribētu izteikt atzinību prof. Ērikai Bizdēnai par ieguldījumu manas zinātniskās darbības pirmsākumos, kas sākās ar viņas vadībā izstrādātu zinātniski pētniecisko darbu Āgenskalna Valsts ģimnāzijas 12. klasē par DNS fragmenta sintēzi, turpinājās ar bakalaura darba izstrādi cukuru ķīmijas jomā un beidzās ar manis "nosūtīšanu" uz LOSI Fizikāli-organiskās ķīmijas laboratoriju. Šis izvērsās par ļoti nozīmīgu pavērsienu manā dzīvē.

Es vēlos pateikties Rīgas Tehniskajai universitātei, Latvijas Organiskās sintēzes institūtam, Valsts Pētījumu programmai un Eiropas Sociālajam fondam par finansiālo atbalstu manu doktorantūras studiju laikā.

Mīļš paldies manai Mammai par ticību maniem spēkiem, rūpēm un atbalstu.

Visbeidzot, es izsaku milzīgu paldies manai mīļotajai Jutai par viņas nenovērtējamo pacietību, sapratni un atbalstu manu doktorantūras studiju laikā un īpaši manas divgadīgās prombūtnes laikā.

GENERAL OVERVIEW OF THE THESIS

Introduction

The recent progress in genome sequencing has generated a vast amount of genomic information that requires a functional interpretation. The three-dimensional structures of proteins (the gene products) are key to understanding their function at a molecular level. In addition to elucidating protein functions, structural information can be used for structure-based drug design. This has contributed to the development of structural genomics, which aims at genome-wide high-throughput protein structure determination and functional annotation. X-ray crystallography and nuclear magnetic resonance (NMR) are the two major methods used in the context of structural genomics for protein structure determination at atomic resolution. Two of the further described protein structures were determined within the structural genomics project "Protein Structure Initiative" while the author was working under supervision of the Nobel Prize laureate, prof. K. Wüthrich at The Scripps Research Institute (USA).

Due to a recent overall increase of biomedical research costs, the industry is seeking to use more rational approaches to solve various medical, biological and biochemical problems. The contribution of target protein structural information to the rational design of new potential drugs and therapeutic tools can be particularly emphasized. NMR spectroscopy is the only biophysical method, capable of atomic level protein structure determination in their natural environment - an aqueous solution. In addition, NMR spectroscopy is well suited for protein-ligand (potential drug) interaction studies and can provide information on low affinity interactions that are not accessible by other methods, as well as establish the interaction sites on the protein surface.

The selected targets are proteins with high biological relevance - DnaI protein participates in DNA replication, R3H domain is found in more than 370 proteins, whose specific function is not always known, while the R3H domain itself might interact with DNA or RNA, carnitine acetyltransferase (CrAT) is involved in energy metabolism, while Nsp7 protein is an essential Severe Acute Respiratory Syndrome (SARS) coronavirus virulence factor.

Aims and objectives

De novo 3D structure determination of proteins DnaI, NP_247299.1 and Nsp7 in solution, prediction and annotation of their molecular function. Analysis of Nsp7 intramolecular rate processes and ns timescale dynamics and evaluation of polymorphisms observed in the previous structures in single crystals and in solution.

Determination of deoxyguanosine monophosphate (dGMP) binding site on the surface of R3H, evaluation of binding thermodynamics of different mononucleotides and creation of a model for R3H ligand selectivity. Determination of the 3D structure of the dGMP-R3H complex.

Characterization of CrAT inhibition by mildronate, modeling of mildronate-CrAT complex.

The following sub-objective was defined for achieving the goals: cloning, bacterial expression and chromatographic purification of ^{13}C , ^{15}N -labeled NP_247299.1 in milligram quantities.

Research methodologies

Protein NMR samples for 3D structure determinations were prepared using recombinant expression in *E. Coli*. The gene fragment coding for target protein was cloned into BL21(DE3) cells using molecular biology techniques. Proteins were expressed in M9 minimal media containing either 1 g/L $^{15}\text{NH}_4\text{Cl}$ and 4 g/L unlabeled D-glucose or 1 g/L $^{15}\text{NH}_4\text{Cl}$ and 4 g/L [$^{13}\text{C}_6$]-D-glucose as the sole sources of nitrogen and carbon. Cells were harvested by centrifugation and lysed by ultrasonication. The soluble fraction of the cell lysate was isolated by centrifugation and purified by affinity chromatography and gel-filtration using ÄKTA Purifier 10 chromatography system. NMR spectra were acquired on Varian Unity Inova 600 MHz, Bruker Avance 600 and 800 MHz spectrometers. The assignment and 3D structure determination of DnaI was based on 3D NOESY- ^{15}N -HSQC, HNHA, 2D NOESY, TOCSY and DQF-COSY NMR spectra. The structure determination of NP_247299.1 employed automated projection spectroscopy for backbone assignment as well as 3D NOESY- ^{15}N -HSQC and two NOESY- ^{13}C -HSQC spectra for automated side chain assignment and collection of NOE distance restraints. The assignments and structure of Nsp7 were determined on basis of triple resonance 3D HNCA, HNCACB, CBCA(CO)NH, HC(C)H-TOCSY spectra and 3D NOESY- ^{15}N -HSQC and NOESY- ^{13}C -HSQC spectra. For characterization of Nsp7 dynamics a 2D $^{15}\text{N}\{^1\text{H}\}$ -NOE spectrum and a series of 2D ^{15}N -HSQC spectra after redissolving a sample lyophilized from $^1\text{H}_2\text{O}$ in $^2\text{H}_2\text{O}$, was recorded. Backbone amide $^1\text{H}/^2\text{H}$ exchange protection factors were determined by fitting the experimental data to an exponential equation. Automated backbone and side chain assignment was performed using the software UNIO-MATCH and UNIO-ATNOS/ASCAN, respectively. Spectra were interactively analyzed using the software XEASY and CARA. Structure calculation was performed using the software CYANA and UNIO-ATNOS/CANDID. Energy minimization in explicit water was performed using CNS and OPALp. Structure validation and analysis was done using the programs MOLMOL, PYMOL, PROCHECK etc.

R3H domain and CrAT ligand interactions were investigated by NMR titration, by monitoring either perturbations of resonance signals or half-width changes in the proton 1D or 2D- ^{15}N -HSQC spectra upon addition of the ligands. Dissociation constants were calculated by fitting of experimental data to a simulated binding isotherm in the program MS Excel. The three-dimensional model of mildronate and CrAT complex was based on saturation transfer difference (STD) NMR spectra, which were used to validate molecular modeling results. Molecular modeling was performed using Molecular Operating Environment software package.

Scientific novelty and main results

- The three-dimensional structure of the helicase-binding domain of a replicative helicase loader DnaI (DnaI-N) has been determined, revealing a novel zinc-binding fold. The chemical shifts and coordinates of the structure have been

deposited in the BioMagResBank (accession code 15926) and the Protein Data Bank (accession code 2K7R).

- The dissociation constants of R3H-mononucleotide complexes have been estimated and the binding site of deoxyguanosine monophosphate (dGMP) on the surface of R3H has been identified, giving functional insights and indications about main interaction partners on the protein. The structure of R3H and dGMP complex has been determined. This has allowed the identification of nucleotide selectivity determining factors of the R3H domain.
- A model for CrAT inhibition by the antiischemic agent mildronate has been established suggesting that the cardioprotective effect of mildronate might be partially mediated by CrAT inhibition.
- A new protocol for X-ray and NMR structure comparison has been developed, and is used to identify local polymorphisms that are found in segments of higher crystallographic B values and correlate with dynamic processes observed in solution.
- The three-dimensional structure of Nsp7 at pH 6.5 has been determined. The study includes characterization of conformational equilibria and intramolecular rate processes in Nsp7, which provide functional insights. It is shown that the residues 70–82 exist in a dynamic equilibrium between α -helix and unstructured conformations that could be required for the formation of the active complex with Nsp8. The chemical shifts of Nsp7 at pH 6.5 were deposited in BioMagResBank under accession number 16981. The atomic coordinates of the 20 conformers representing the solution structure of Nsp7 at pH 6.5 were deposited in the Protein Data Bank under accession code 2KYS.
- The application of an automated protein structure determination protocol used by the Joint Center for Structural Genomics (JCSG) is demonstrated by 3D structure determination of NP_247299.1. The chemical shifts of NP_247299.1 have been deposited in BioMagResBank (accession No. 16389) and the atomic coordinates of the 20 conformers representing the NMR structure were deposited in the Protein Data Bank (accession code 2KLA).

Applications

The determined 3D protein structures and data on protein-ligand interactions can be used for structure-based drug design. The experimental protein structures, particularly those showing new and previously unknown folds, can be used as templates for structure modeling of other proteins on basis of amino acid sequence similarity.

Thesis structure and length

The thesis is presented as a set of scientific publications on the use of NMR spectroscopy in protein studies. It consists of 5 publications in peer-reviewed journals that are included in internationally accessible databases (Science Citation Index and Chemical Abstracts). The publications are written in English and constitute a total of 47 pages.

Publications and approbation of the thesis

The thesis findings are set out and discussed in 5 publications:

1. Loscha K.V., Jaudzems K., Ioannou C., Su X.-C., Hill F.R., Otting G., Dixon N.E., Liepinsh E. A novel zinc-binding fold in the helicase interaction domain of the *Bacillus subtilis* DnaI helicase loader// *Nucleic Acids Research*. - 2009. - Vol. 37. - p. 2395-2404.
2. Jaudzems K., Zhulyenkov D., Liepinsh E. R3H domain complexes with mononucleotides// *Scientific Journal of RTU. Material Science and Applied Chemistry*. - 2009. - Vol. 20. - p. 58-63.
3. Jaudzems K., Kuka J., Gutsaits A., Zinovjevs K., Kalvinsh I., Liepinsh E., Liepinsh E., Dambrova M. Inhibition of carnitine acetyltransferase by mildronate, a regulator of energy metabolism// *Journal of Enzyme Inhibition and Medicinal Chemistry*. - 2009. - Vol. 24. - p. 1269-1275.
4. Jaudzems K., Geralt M., Serrano P., Mohanty B., Horst R., Pedrini B., Elsliger M.-A., Wilson I.A., Wüthrich K. NMR structure of the protein NP_247299.1: comparison with the crystal structure// *Acta Crystallographica Section F*. - 2010. - Vol. 66. - p. 1367-1380.
5. Johnson M.A., Jaudzems K., Wüthrich K. NMR structure of the SARS coronavirus nonstructural protein Nsp7 in solution at pH 6.5// *Journal of Molecular Biology*. - 2010. - Vol. 402. - p. 619-628.

The thesis results have been presented at the following conferences and scientific seminars:

1. Jaudzems K., Zhulyenkov D., Liepinsh E. R3H domain complexes with mononucleotides. Poster at the RTU 49th International Scientific Conference. Riga, October 13, 2008.
2. Jaudzems K. Automation in the NMR structure determination of a 206-residue protein. Oral presentation at the Joint Center for Structural Genomics Annual Meeting. San Diego, USA, April 23, 2009.
3. Jaudzems K., Mohanty B., Serrano P., Geralt M., Liepinsh E., Wüthrich K. Solution NMR structure of the protein NP_247299.1: Comparison with the crystal structure. Poster at the RTU 50th International Scientific Conference. Riga, October 16, 2009.
4. Jaudzems K. NMR based structural genomics within the framework of the "Protein Structure Initiative" (in Latvian). Seminar at the Latvian Institute of Organic Synthesis. Riga, January 6, 2010.
5. Jaudzems K., Geralt M., Serrano P., Mohanty B., Horst R., Pedrini B., Elsliger M.-A., Wilson I.A., Wüthrich K. NMR structure of the protein NP_247299.1: comparison with the crystal structure. Poster at the Joint EUROMAR 2010 and 17th ISMAR Conference. Florence, Italy, July 4-9, 2010.
6. Jaudzems K., Zhulyenkov D., Otting G., Liepinsh E. Structural and thermodynamic characterization of nucleotide binding to the R3H domain from human Subp-2. Poster at the 18th International Conference of Methods in Protein Structure Analysis. Uppsala, Sweden, August 25-28, 2010.
7. Jaudzems K. Nucleotide binding to the R3H domain studied by NMR. Oral presentation at the 1st EAST-NMR Young Investigators Meeting. Istanbul, September 29, 2010.

CONTENTS OF THE THESIS

The thesis contains 5 scientific publications on a common subject – the use of NMR spectroscopy for studies of protein structure, function and ligand interactions. The present doctoral thesis summary is divided into two chapters. The first chapter describes protein structure determinations by NMR in solution and the second chapter describes protein-ligand interaction studies. The chapters are further divided into sections, devoted to a concise outline of each scientific publication, including a brief introduction, main results and conclusions. Each publication describes the use of a different set of NMR techniques altogether forming an almost complete description of the available protein NMR methods.

CHAPTER 1. PROTEIN STRUCTURE DETERMINATION BY NMR

The determination of 3D protein structure by NMR involves several successive steps: preparation of a mM concentration protein sample, recording of NMR spectra, NMR resonance assignment (connecting the NMR signals to distinct atoms or groups of atoms in the polypeptide), structural interpretation of NMR signals (conformational restraint collection) and at last, 3D structure calculation and refinement in explicit solvent. The first three steps are the most effort- and time-consuming.

The choice of NMR experimental and assignment strategies depends on the type of protein labeling used - the options are: no labeling, ^{15}N or double ^{15}N and ^{13}C -labeling. The uniform labeling with naturally low abundant NMR-active nuclei permits the use of higher dimensionality experiments that significantly improve spectral resolution and facilitate assignment, which is crucial for high quality structure determination. In this work, strategies for the analysis of ^{15}N (DnaI example) and doubly ^{15}N , ^{13}C -labeled proteins (examples of NP_247299.1 and Nsp7) are described. The NP_247299.1 paper describes the use of automated projection spectroscopy involving very high dimensionality experiments (4D and 5D) as well as partial automation of resonance assignment to save time.

In addition to 3D structure determination, NMR spectroscopy can give insights into protein dynamics, structural stability and conformational equilibria that exist in solution. The use of such techniques is shown in the papers describing the structure determinations of NP_247299.1 and Nsp7. In the work of NP_247299.1 a new protocol for NMR and crystal structure comparison is also described.

1.1. A NOVEL ZINC-BINDING FOLD IN THE HELICASE INTERACTION DOMAIN OF THE *B. SUBTILIS* DNAI HELICASE LOADER

DNA replication in all organisms is carried out by replisomes, multiprotein machines that contain a ring-shaped DNA helicase for separation of the strands of the DNA duplex [1]. In bacteria, the helicase forms a hexameric structure that encircles single-stranded (ss) DNA. In *B. subtilis* this process is assisted by the helicase loader protein DnaI and a pair of DNA-remodelling co-loader proteins (called DnaB and DnaD), which guide the helicase-DnaI complex to specific sites in the DNA [2]. However, *in vitro* and in the presence of ATP, DnaI alone is sufficient for loading of the helicase onto ssDNA [3-4]. DnaI consists of two structured domains. The larger C-terminal domain contains the ATP-binding site and a cryptic site for ssDNA binding,

whereas the N-terminal domain (DnaI-N) seems to be primarily responsible for helicase binding [4]. Here we present the NMR structure of the helicase-binding domain of a replicative helicase loader, the N-terminal domain of DnaI, and identify its zinc-coordinating residues.

NMR spectroscopy cannot be used to determine directly the location of the zinc ion, so we confirmed the identity of zinc ligands by independent methods. DnaI-N contains one histidine, four cysteines and one zinc ion [4]. The involvement of His84 in zinc binding was probed using a ^{15}N -HSQC spectrum with the INEPT delays set to 20 ms yielding correlations via $^2J_{\text{HN}}$ couplings of the histidine side chain. The ^{15}N chemical shift of the $\text{N}^{\epsilon 2}$ resonance of His84 (221 ppm) [5] was characteristic of a zinc-bound imidazole nitrogen, confirming that the side chain of His84 is one of the zinc ligands. In addition, there was no sign of structural heterogeneity, indicating that the zinc ion was present in all protein molecules and in the same coordination environment. The zinc-coordinating cysteines were identified by thiol titrations. DnaI-N gave biphasic pseudo first-order kinetics, with one exposed thiol group titrating quickly, and three others about 50-fold more slowly. In contrast, titration of the DnaI-Cys76Ala mutant protein showed only single-phase kinetics, with three thiols titrating slowly, thus indicating that Cys76 is not involved in zinc coordination.

The structure of DnaI-N comprises four α helices and two β strands as well as a short 3_{10} helix near the C-terminus (Fig. 1A and C). The structure presents a novel fold, as a search of the Protein Data Bank using the program Dali [6] failed to reveal a domain of similar structure. The zinc ion is coordinated by Cys67, Cys70, His84 and Cys101 (Fig. 1A). The zinc ion probably plays an important structural role, as it ties together three sequentially distant segments of the polypeptide chain in a region, where the structure has few hydrophobic residues (Fig. 1D). The N-terminal 14 residues were disordered. As expected for a mobile polypeptide segment, the NMR signals of the N-terminal residues were also significantly narrower than those of the structurally defined part of the protein. Remarkably, the positions of Ile4, Leu8 and Val11 in the unstructured N-terminus are also occupied by hydrophobic residues in other organisms, suggesting a functional role for these flexible and highly solvent accessible residues, presumably in protein–protein interactions. Similarly, the ^1H NMR signals of residues 63–82 showed a reduced line width. Conceivably, these loop residues following helix 4 could assume a single, rigid conformation in the presence of the C-terminal ATP-binding domain. A longer construct of DnaI-N123 showed significant chemical shift changes for the loop residues suggesting an interaction with the segment following the C-terminus of DnaI-N (Fig. 1B). This interpretation would be consistent with the observation that the NMR signals of residues 107–123 were particularly broad yet were at chemical shifts indicative of random coil structures.

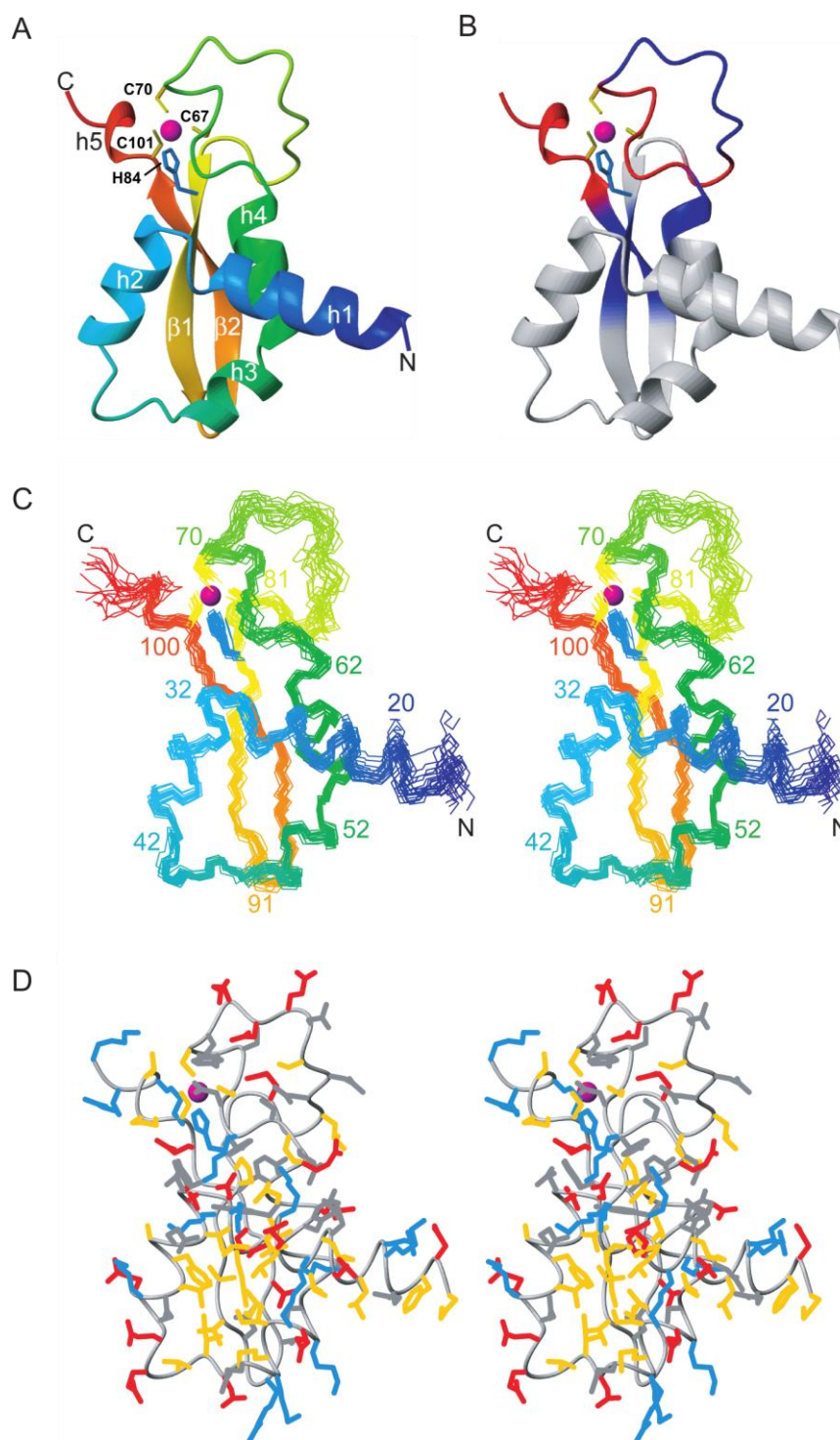


Fig. 1. Solution structure of DnaI-N. Only the ordered residues 15–106 are shown. The zinc atom is shown as a sphere (magenta) and the side chains of coordinating residues Cys67, Cys70, His84 and Cys101 are shown as sticks.

- (A) Ribbon representation of DnaI-N. The secondary structure elements are identified. (B) Same as (A), but color coded to reflect the changes in backbone amide chemical shifts observed in DnaI-N vs. those in DnaI-N123 (blue: small changes >0.015 ppm (^1H) or >0.15 ppm (^{15}N); red: significant changes >0.05 ppm (^1H) or >0.5 ppm (^{15}N)). (C) Stereo view of a superposition of the backbone atoms of 20 NMR conformers. (D) Stereo view of a heavy-atom representation with sticks. The side-chains are color coded in blue (Lys, Arg, His), red (Asp, Glu), yellow (Ala, Cys, Ile, Leu, Met, Phe, Pro, Trp, Val) and gray (Asn, Gln, Ser, Thr, Tyr)

1.2. NMR STRUCTURE OF THE PROTEIN NP_247299.1: COMPARISON WITH THE CRYSTAL STRUCTURE

The hypothetical protein NP_247299.1 is the gene product of locus MJ0327 in the genomic sequence of the *Methanococcus jannaschii* genome. In this paper, we describe the NMR solution structure of NP_247299.1 obtained using a new protocol for automated protein structure determination that was implemented by the JCSG NMR Core, and present a systematic comparison of the results obtained by largely automated solution structure determination with the crystal structure (PDB code 2qtd).

The automation of the NMR structure determination of NP_247299.1 provided the following results: the polypeptide backbone resonance assignments were obtained from the automated projection spectroscopy (APSY; [7-8]) NMR experiments using the software UNIO-MATCH [9], which yielded assignments for the $^1\text{H}^{\text{N}}$, ^{15}N , $^{13}\text{C}^{\alpha}$ and $^1\text{H}^{\alpha}$ atoms of 97 residues (94%) and for the $^{13}\text{C}^{\beta}$ atoms of 91 residues of all 104 residues. The backbone and $^{13}\text{C}^{\beta}$ assignments were validated and extended interactively, and followed by automated side-chain assignment with the program UNIO-ATNOS/ASCAN [10]. Interactive inspection showed that about 95% of these assignments were correct and, for most of the side chains with partial assignment, the chemical shift lists could be expanded interactively. Structure calculation was performed using UNIO-ATNOS/CANDID [11-12] and CYANA [13], incorporating automated restraint collection and NOESY peak assignment. The total time used for the interactive steps was about 70 h, which is significantly less time than in the case of fully interactive structure determination (then it could take several months).

The NMR structure of NP_247299.1 comprises three α -helices, five β -strands and a short 3_{10} -helix in the sequential order $\beta 1$ – 3_{10} – $\beta 2$ – $\beta 3$ – $\alpha 1$ – $\beta 4$ – $\alpha 2$ – $\beta 5$ – $\alpha 3$ (Fig. 2A). The β -strands form a twisted sheet with topology 3–2–1–4–5, in which strands $\beta 3$, $\beta 1$, $\beta 4$ and $\beta 5$ are parallel to each other and strand $\beta 2$ is antiparallel. Helices $\alpha 1$ and $\alpha 2$ are arranged in parallel at a small angle, with the 3_{10} -helix positioned between them at their N-terminal ends. This combination of three helices is docked against one side of the twisted β -sheet, whereas helix $\alpha 3$ is located on the opposite side. The high global structure similarity between the NMR and X-ray structures of NP_247299.1 is visualized by superposition of the crystal structure and the bundle of NMR conformers (Fig. 2B).

To assess the accuracy and precision of the result from the automated approach, we performed a detailed comparison with the crystal structure. The identification of locally variable sites was guided by a search for sequence locations with high B values in the crystal structure or/and high variation within the bundle of 20 NMR conformers (Fig. 3), as well as locations of variable dihedral angles or different packing. Characterization of these local ‘hot spots’ was then supported by additional NMR data that are differently averaged in a dynamic conformational ensemble than the ^1H – ^1H NOEs, i.e. scalar $^1\text{H}^{\text{N}}$ – $^1\text{H}^{\alpha}$ spin–spin couplings, $^3J_{\text{HN}\alpha}$. This allowed to draw the following conclusions: the protein core residues with solvent accessibility below 15% assume a nearly identical conformation in the crystal and solution structures. Locally increased crystallographic B values correlate with dynamic structural polymorphisms in solution, including that the solution state of the protein involves a slow dynamic equilibrium on a time scale of milliseconds or slower between two ensembles of rapidly interchanging conformers that contain, respectively, the cis or trans form of the C-terminal proline and represent about 25 and 75% of the total protein (Fig. 4).

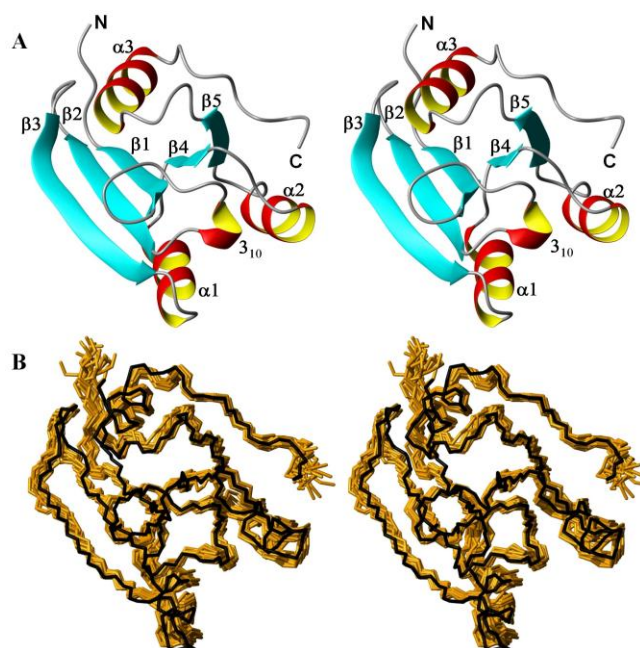


Fig. 2. NMR structure of protein NP_247299.1 and comparison of the NMR structure with the crystal structure.

- (A) Stereo ribbon diagram. The secondary structure elements are identified.
 (B) Stereo view of a superposition of the polypeptide backbone heavy atoms of the crystal structure (black line) with the 20 NMR conformers (brown)

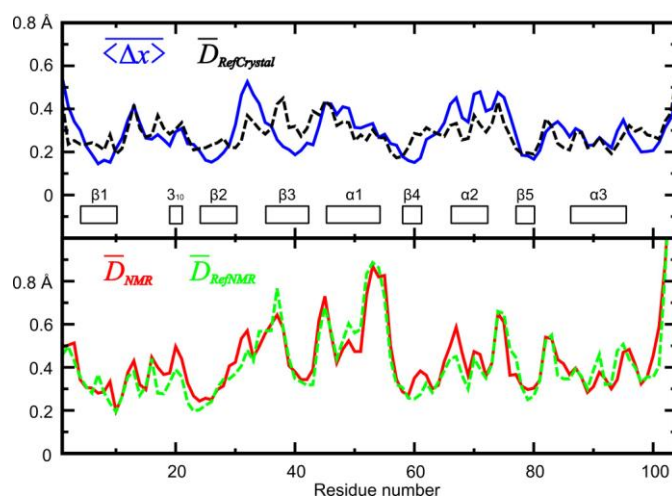


Fig. 3. Comparison of crystallographic B values and per-residue NMR displacements. Upper panel represents data for crystal structure and reference crystal structure. Lower panel represents data for NMR structure and reference NMR structure. For the crystal structure, per-residue displacements were calculated from the B values with the equation $\overline{\langle \Delta x \rangle} = c \cdot \overline{B}$, where the coefficient c was determined by linear regression analysis

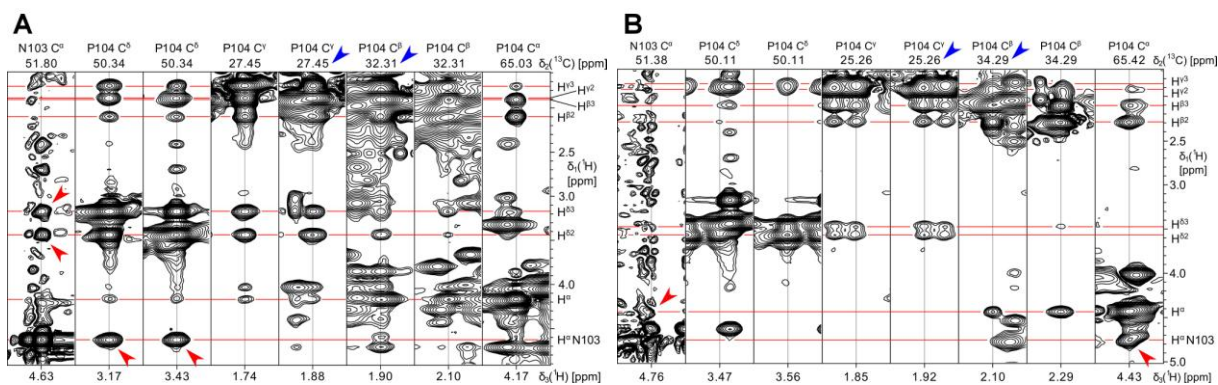


Fig. 4. Cis–trans isomerization of Pro104. Strips from a 3D ^{13}C -resolved $[\text{}^1\text{H}, \text{}^1\text{H}]$ -NOESY spectrum representing the signals of the trans and cis forms of the Asn103–Pro104 peptide bond.

- (A) The trans form of Pro104 is manifested by $d_{\alpha\delta}^{NP}$ NOE connectivities (red arrows) and by the typical $^{13}\text{C}^{\beta}$ and $^{13}\text{C}^{\gamma}$ chemical shift pattern (blue arrows).
- (B) The cis form of Pro104 is manifested by $d_{\alpha\alpha}^{NP}$ NOE peaks (red arrows) and the large difference of about 9 p.p.m. between the $^{13}\text{C}^{\beta}$ and $^{13}\text{C}^{\gamma}$ chemical shifts (blue arrows).

1.3. NMR STRUCTURE OF THE SARS-COV NONSTRUCTURAL PROTEIN 7 IN SOLUTION AT pH 6.5

The severe acute respiratory syndrome coronavirus (SARS-CoV) nonstructural protein Nsp7 is of interest for its potential roles in the transcription and replication of the viral RNA genome. An important role of Nsp7 is also indicated by the observation that deletion of Nsp7 or mutation of the Nsp7/Nsp8 proteolytic cleavage site is lethal to murine hepatitis virus [14]. The NMR structure of Nsp7 has been previously determined in solution at pH 7.5 [15], and crystal structure determination was reported for a hexadecamer complex with Nsp8 [16], which has been shown to have RNA primase activity. [17]. The complex in the crystal contains eight molecules each of Nsp7 and Nsp8, forming a channel with appropriate width and charge to accommodate double-stranded RNA. [16]. In both of these structures, the polypeptide fold of Nsp7 includes four helices as the only regular secondary structures, which cover approximately 60% of the 85-residue polypeptide chain. A comparison of the Nsp7 folds in the two structures revealed extensive differences in the sequence positions and lengths of the four helices, as well as in the spatial arrangement of the helices in the tertiary structure. In view of the implicated conformational polymorphism, we decided to further investigate the behavior of Nsp7 in different environments.

The present project was initiated by the screening of a wide range of solution conditions in search of protein samples that would enable the recording of high-quality NMR data and the collection of more extensive sets of conformational constraints than had been possible at pH 7.5 and high ionic strengths. Based on the results of this screening, an aqueous solution at pH 6.5 with low ionic strength was selected for a new NMR structure determination of Nsp7. $^{13}\text{C}^{\alpha}$ chemical shifts showed that Nsp7 at pH 6.5 is an α -protein with four helices (Fig. 5A) and structure calculation resulted in a high-quality NMR structure containing an antiparallel bundle of four helices (Fig. 6). The polypeptide segment 1–10 forms a disordered tail, helix $\alpha 1$ (residues 13–20) is

linked by a well-defined loop of residues 21–28 with helix $\alpha 2$ (residues 29–42). A short segment of nonregular secondary structure leads to helix $\alpha 3$ (residues 47–65), and a short loop connects to helix $\alpha 4$ (residues 70–82). The protein was further characterized by a steady-state $^{15}\text{N}\{^1\text{H}\}$ -NOE experiment, which is sensitive to the picosecond to nanosecond timescale mobility of the polypeptide backbone. Increased subnanosecond mobility was evident for the N-terminal decapeptide segment and the C-terminal hexapeptide, which includes the last turn of helix $\alpha 4$ (Fig. 5B), whereas the data for residues 13–79 show that this central polypeptide segment forms a compact globular fold.

To gain insights into Nsp7 conformational equilibria in solution at pH 6.5, we measured the backbone amide $^1\text{H}/^2\text{H}$ exchange rates (Fig. 5C, expressed as amide proton protection factors (Pf) [18]) and performed line shape analysis. The protection factors reflect the degree of protection of each proton by protein secondary and tertiary structures. Protons that are involved in the hydrogen bonds of regular secondary structures or that are otherwise sequestered from the solvent experience lower rates of chemical exchange (higher protection factors). In Nsp7, helices $\alpha 2$ and $\alpha 3$ have the highest protection factors, followed by $\alpha 1$ and $\alpha 4$; for polypeptide segments with nonregular secondary structure, the protection was too small to be measured with the standard approach used here (Fig. 5C). Of special interest is the behavior of helix $\alpha 4$. The reduced $\Delta\delta(^{13}\text{C}^\alpha)$ values of residues 78–82 (Fig. 5A), and the comparatively low protection factors for the entire helix indicate that there is a reduced population of $\alpha 4$ in the NMR structure because of a dynamic equilibrium with unstructured solvent accessible conformations. Since sub-nanosecond timescale motion of the protein backbone was observed only for the C-terminal hexapeptide segment of the protein (Fig. 5B), there was an indication that the conformational equilibria involving helix $\alpha 4$ are governed by slower motions. This indication was confirmed by a line shape analysis - the residues Ser63 and Gln65 near the C-terminal end of $\alpha 3$, Val68 in the loop joining $\alpha 3$ and $\alpha 4$, and Asn71 and Leu73 at the start of $\alpha 4$ all exhibit a pronounced line broadening indicative of ms timescale motions.

From the present study we conclude that helices $\alpha 2$ and $\alpha 3$ form a conserved core of the Nsp7 structure, with their lengths, positions, and relative orientations being largely preserved in different environments, and with helices $\alpha 1$ and $\alpha 4$ adopting quite different lengths, positions in the sequence, and relative orientations (Fig. 7). The crystal structure of the complex with Nsp8 differs from the pH 6.5 solution structure primarily by a rotation of helix $\alpha 4$ away from the $\alpha 2/\alpha 3$ core and by helical folding of the polypeptide segment of residues 3–12 (Fig. 7A). The long helix $\alpha 1$ in the crystal structure (Fig. 7C), which would not be compatible with the amide proton protection factors measured in solution at pH 6.5 (Fig. 5C), then occupies the position taken by $\alpha 4$ in the pH 6.5 solution structure (Fig. 7A). In the pH 7.5 solution structure, $\alpha 4$ is packed against $\alpha 3$ and has no contacts with $\alpha 2$, so that helices $\alpha 2$, $\alpha 3$, and $\alpha 4$ line up to form a flat three-helix sheet (Fig. 7B), with $\alpha 1$ packed at an angle of about 45° against $\alpha 2$ and $\alpha 3$. This suggests that Nsp8 binds to a transient form of Nsp7, in which helix $\alpha 4$ is unfolded (Fig. 5A and C). Binding to Nsp8 would trigger the formation of the short helix $\alpha 4$ immediately after $\alpha 3$, as well as the formation of the long helix $\alpha 1$ (Fig. 7C), which then takes up the position occupied by $\alpha 4$ in the pH 6.5 solution structure (Fig. 7A). In the crystal structure, the locations of $\alpha 1$ and $\alpha 4$ are stabilized by numerous contacts with hydrophobic side chains of Nsp8. Interestingly, while the Nsp7 sequence in helices $\alpha 1$ – $\alpha 3$ is highly conserved among all CoVs, the sequence of

helix $\alpha 4$ is quite variable. These findings indicate that locking of helix $\alpha 4$ of Nsp7 in a conformation that prevents binding to Nsp8 could impair the replication machinery of the virus, which might provide the lead for future drug designs.

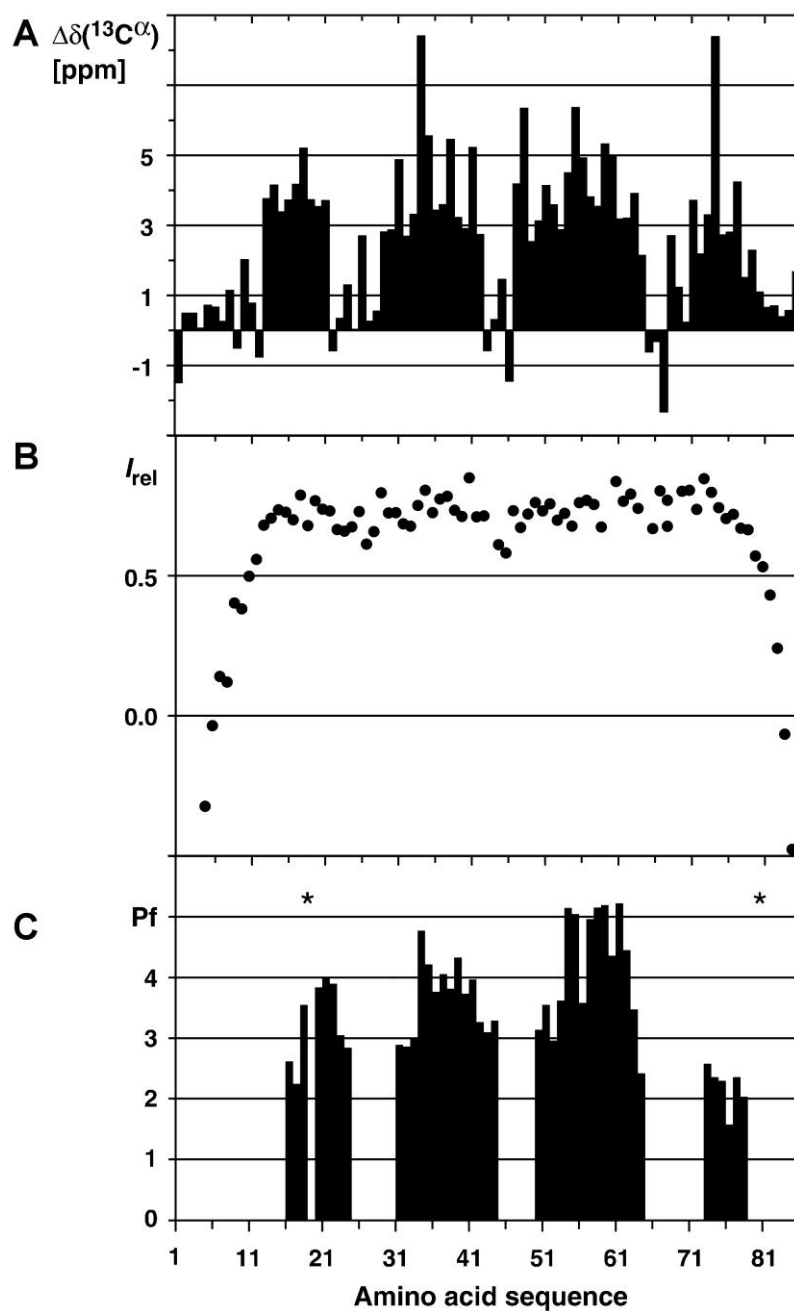


Fig. 5. Plots of NMR data measured in solution at pH 6.5 versus the Nsp7 sequence.
 (A) $^{13}\text{C}^\alpha$ chemical shift deviations from random-coil values, $\Delta\delta(^{13}\text{C}^\alpha)$.
 (B) $^{15}\text{N}\{^1\text{H}\}$ -NOE values, I_{rel} .
 (C) Amide $^1\text{H}/^2\text{H}$ exchange protection factors. Asterisks indicate positions where no measurements could be made due to spectral overlap in the $[^{15}\text{N}, ^1\text{H}]$ -HSQC spectrum.

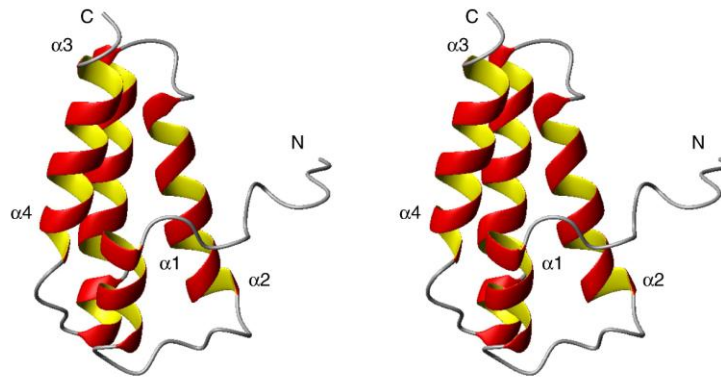


Fig. 6. Stereo ribbon representation of Nsp7 structure in solution at pH 6.5. The secondary structure elements are identified.

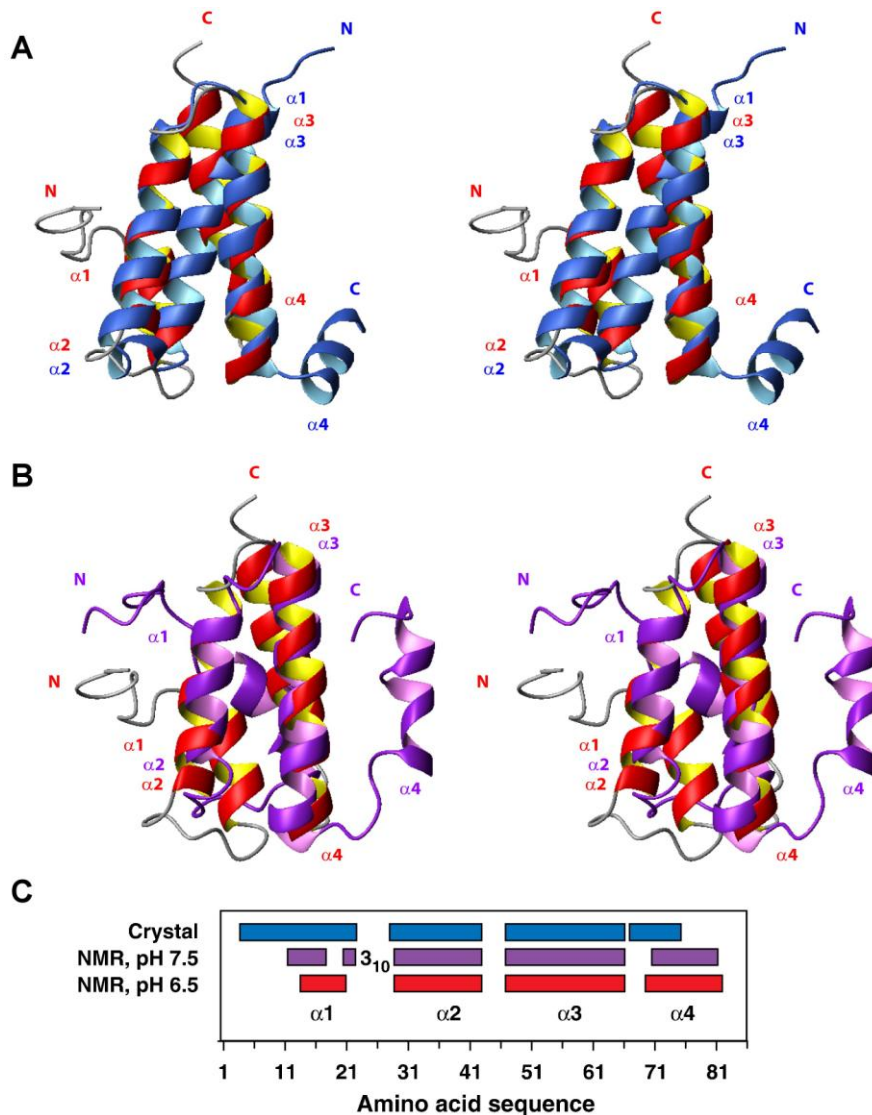


Fig. 7. Stereo ribbon drawings of the superpositions of the solution structure of Nsp7 at pH 6.5 (red/yellow) with the two previously reported Nsp7 structures. The structures were superimposed for minimal RMSD of the backbone of helices $\alpha 2$ and $\alpha 3$.

(A) Superposition with the crystal structure of Nsp7 in a complex with Nsp8 (blue).

(B) Superposition with the solution structure of Nsp7 at pH 7.5 (purple)

(C) Locations of α -helices in all the experimentally determined Nsp7 structures

CHAPTER 2. NMR STUDIES OF PROTEIN-LIGAND INTERACTIONS

NMR spectroscopy is a powerful method for detecting and quantifying binding of small molecules to proteins. It can be used for screening of large compound libraries, measurement of complex dissociation constants, determination of ligand binding site on the protein surface or even three-dimensional structure determination of a protein-ligand complex.

The NMR experiments used for screening purposes and protein-ligand interaction studies can be divided into protein-based and ligand-based methods. Besides detection of binding, the protein-based methods can provide information on the interacting protein residues, although they require a relatively large amount of protein (^{15}N -labeled) per experiment. These methods rely on ligand-induced local magnetic environment changes around the protein atoms near to the ligand binding site, observable in the spectrum as shifting of protein signals. The use of protein-based methods is described in the R3H domain and mononucleotide complexes' paper. The use of the ligand-based Saturation Transfer Difference (STD) method is described in the CrAT studies. This method is based on the direct intermolecular magnetization transfer between the protein and a bound ligand. The ligand-based methods require a relatively small amount of unlabeled protein allowing the screening of larger compound libraries, however the information on protein interaction sites is lacking.

2.1. R3H DOMAIN COMPLEXES WITH MONONUCLEOTIDES

The R3H domain has previously been identified as a conserved sequence motif in putative nucleic-acid-binding proteins from diverse range of organisms [19]. It is named after the presence of an invariant arginine residue and a highly conserved histidine residue that are separated by three residues. The 3D solution structure of the R3H domain from human *Subp-2* has been determined by NMR spectroscopy and shown to represent the R3H domains in general [20]. Here, we report our results from NMR binding studies of the R3H domain and mononucleotides, aimed at establishing the ligand specificity and selectivity determining factors of the R3H domain.

Several NMR techniques can be used to detect and quantify the binding of small molecules to proteins [21]. We chose to observe the ^1H NMR chemical shifts of the protein signals in the high-field-shifted methyl group region as a function of small molecule concentration. For a molecule in fast exchange between the bound and free forms any NMR parameter is the mole-fraction weighted average of that belonging to the two different states. Thus, by acquiring a series of experiments and applying non-linear least-squares fitting of simulated data to experimental data, the complex dissociation constant K_D could be determined. Results from the titration of the R3H domain with guanosine mono-phosphate (GMP) are shown in Fig. 8(A). The relationship between $\delta_{(\text{obs})}$ and GMP concentration for the singlet at 0.26 ppm along with the fitted binding curve that determines the K_D , is shown in Fig. 8(B). The results of the NMR chemical shift titrations of the R3H domain with mononucleotides are summarized in Table 1.

To identify the putative DNA binding site on the R3H structure, we monitored the backbone amide chemical shift changes in the 2D [^1H , ^{15}N]-HSQC spectrum upon addition of dGMP. Most of the shifting signals were found to belong to residues of helix $\alpha 2$ and the strands $\beta 2$ and $\beta 3$ that are located on the same face of the R3H

domain and include the R3H motif, thereby verifying its role in nucleic acid recognition. On the basis of the obtained results a model of R3H and dGMP interactions was proposed. The mode of interaction was recently verified by 3D structure determination of the dGMP-R3H complex (Fig. 9).

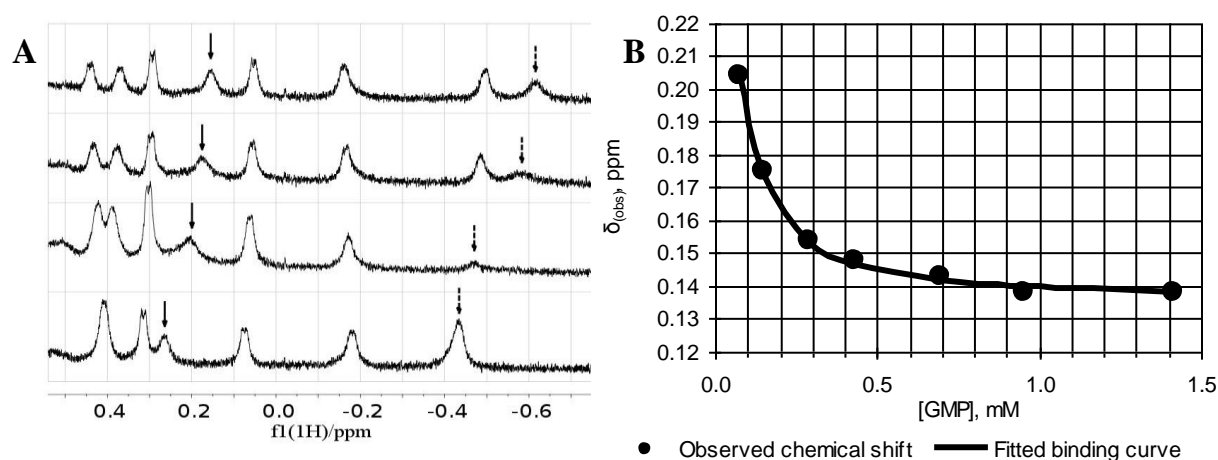


Fig. 8. (A) Methyl-group region of the proton 1D NMR spectra of R3H at different R3H to GMP concentration ratios. The arrows indicate the signals that were used for K_D determination.

(B) The chemical shift of the signal at 0.26 ppm as a function of GMP concentration (mM) at constant R3H concentration (0.1 mM). The circles are the experimental points, and the line is the fitted binding isotherm

Table 1

Dissociation constants of the R3H domain complexes with mononucleotides

N ^o	Nucleotide (abbreviation)	K_D , μ M	N ^o	Nucleotide (abbreviation)	K_D , μ M
1	GMP	33.6 ± 1.0	5	dGMP	8.0 ± 0.1
2	AMP	109.0 ± 4.0	6	dAMP	58.7 ± 4.7
3	UMP	106.4 ± 10.6	7	dTMP	268.5 ± 54.5
4	CMP	290.5 ± 30.5	8	dCMP	276.0 ± 43.0

The obtained results show that the R3H domain binds mononucleotide mono-phosphates, and the dissociation constants of the complexes are within the micromolar region. It has a substantial preference for binding with deoxy-ribonucleotides over ribonucleotides, especially for the purine-containing bases. The R3H domain does not bind mononucleotide di-phosphates and tri-phosphates as well as un-phosphorylated mononucleotides not present in the DNA and the RNA. The R3H domain has no affinity towards 3'-phosphorylated mononucleotides too. The binding site of deoxy-guanosine mono-phosphate (dGMP) is identified from chemical shift perturbations in the 2D [1 H, 15 N] HSQC spectrum and later verified by 3D structure determination of the dGMP-R3H complex. Our data show that binding of dGMP occurs through electrostatic interactions with the positively charged sidechains of Arg47 and Arg70, π - π stacking interactions with His51 and several hydrogen bonds (Fig. 9B). His51 could be responsible for selectivity against purine-containing nucleotides that provide better stacking interactions and the hydrophobic side chain of Leu48 could be responsible for selectivity against deoxy-ribonucleotides.

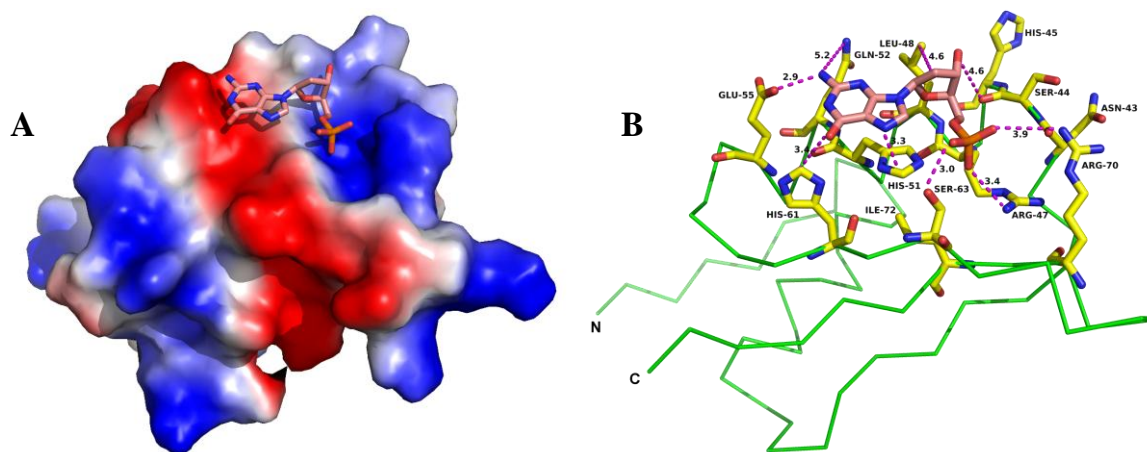


Fig. 9. Solution structure of the R3H domain in complex with dGMP. (A) dGMP binding site on the molecular surface of the R3H domain. The surface is colored by electrostatic potential with negative potential in red and positive potential in blue. (B) Zoom-in on the dGMP (pink) binding site. All residues within 5 Å of the ligand are identified and shown as yellow sticks. The nitrogen and oxygen atoms are colored in blue and red, respectively. Hydrogens are omitted for clarity. The potential interactions are drawn with dashed lines and the distances between interaction partners are indicated

2.2. INHIBITION OF CARNITINE ACETYLTRANSFERASE BY MILDRONATE, A REGULATOR OF ENERGY METABOLISM

Carnitine acetyltransferase (CrAT) catalyzes the reversible transfer of acetyl groups between acetyl-coenzyme A (acetyl-CoA) and L-carnitine [22], thus regulating the cellular pool of CoA and the availability of activated acetyl groups. [23]. Mildronate (3-(2,2,2-trimethylhydrazinium)propionate) is a cardioprotective drug that acts as a L-carnitine concentration lowering agent and inhibitor of free fatty acid β -oxidation [24-25]. Structurally, mildronate is related to L-carnitine and its bioprecursor, γ -butyrobetaine (GBB). In this study, biochemical measurements, saturation transfer difference (STD) nuclear magnetic resonance (NMR) spectroscopy, and molecular docking were applied to give insights into the CrAT binding of a synthetic inhibitor, the cardioprotective drug mildronate.

For quantification of the binding affinity of mildronate towards CrAT, an NMR titration was carried out. The linewidths of the ^1H NMR signals, an approximation to the relaxation rate constant $1/T_2$, of mildronate were monitored as a function of the changing solution composition. Due to very small linewidth changes during titration (only about 0.5 Hz) for the signals chosen and the inherent sensitivity of the fitting procedure to small errors in the measured linewidth, normalization to an internal reference (DSS) was used to compensate for magnetic field inhomogeneity. The experimental data points were fitted to a simulated binding isotherm by changing the two parameters, complex dissociation constant (K_D) and linewidth of the bound form (ν_{bound}). The binding isotherms for the proton signals of mildronate are shown in Fig. 10. The estimated K_D was $101 \pm 19 \mu\text{M}$.

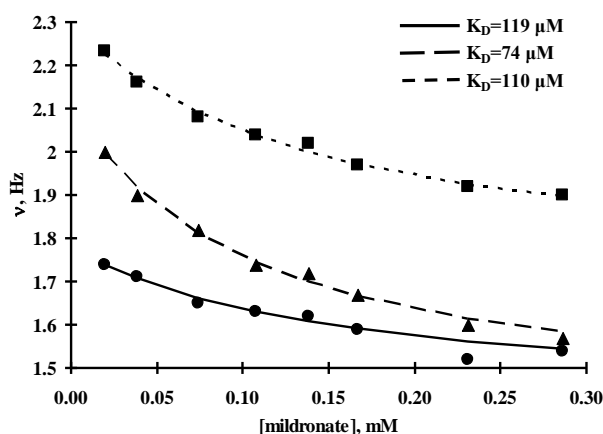


Fig. 10. The linewidth of trimethylammonium group (circles), α -methylene group (triangles), and β -methylene group (squares) proton signals of mildronate as a function of mildronate concentration (mM) at constant CrAT concentration (20 μ M). The circles, squares, and triangles are the experimental points, and the curves are the fitted binding isotherms.

The mildronate binding site on CrAT was identified using competition STD-NMR experiments [26]. In the case of two competing ligands, addition of the second ligand to a sample of an indicator ligand and the protein causes disappearance or reduction of the STD signals of the indicator ligand. The competing ligand displaces the indicator ligand in its binding site on the protein; therefore, less magnetization from the bound state of the indicator ligand can be transferred into solution for detection. The extent of signal reduction of the indicator ligand is subject to the binding affinity of the competing ligand. The gradual addition of mildronate to the carnitine–CrAT complex caused a significant and continuous reduction of the STD signals of carnitine and an increase of the mildronate signals (Fig. 11A) indicating that mildronate binds to the same binding site as carnitine and suggesting a mildronate binding constant value in the mM range, similar to that of carnitine. Furthermore, the same principle was used to prove that mildronate does not compete for binding with acetyl-CoA (Fig. 11B).

The saturation transfer difference NMR spectra contain information on the ligand groups that come closer to the protein surface in the ligand-bound state and this is reflected in the STD amplification factors. The STD amplification factors of carnitine α -CH₂, γ -CH₂ and trimethylammonium signals (17%, 8.5% and 37%, respectively) are in good agreement with the carnitine conformation in the CrAT-carnitine complex crystal structure (PDB code 1S5O): the γ -CH₂ group of carnitine is exposed to the solvent, resulting in the smallest STD amplification factor; meanwhile, the trimethylammonium group is buried in a hydrophobic pocket, giving close contacts to the surface of the protein and resulting in a higher STD amplification factor. The STD amplification factors of mildronate α -CH₂, β -CH₂ and trimethylammonium signals (11%, 17% and 1% respectively) suggested a bound conformation with a different position of the trimethylammonium group, possibly exposed to the solvent. This indication was confirmed by molecular docking studies (Fig. 12).

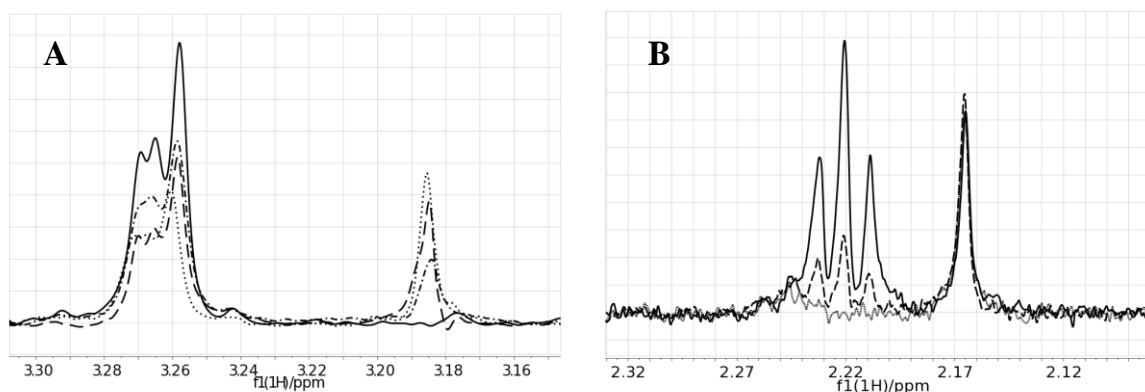


Fig. 11. (A) Titration of the competing ligand mildronate to a sample of carnitine and CrAT. Series of 1D STD-NMR spectra showing the γ -methylene signal of carnitine at 3.265 ppm and the trimethylammonium group signal of mildronate at 3.185 ppm (B) Titration of the non-competing ligand mildronate to a sample of acetyl-CoA and CrAT. Series of 1D STD-NMR spectra showing the α -methylene signal of mildronate at 2.22 ppm and a signal of acetyl-CoA at 2.165 ppm

The obtained results show that mildronate inhibits CrAT in a competitive manner through binding to the carnitine binding site and not the acetyl-CoA site. The bound conformation of mildronate closely resembles that of carnitine except for the orientation of the trimethylammonium group, which in the mildronate molecule is exposed to the solvent. The dissociation constant of mildronate and CrAT complex is approximately 0.1 mM, and the K_i is 1.63 mM. The obtained results suggest that the cardioprotective effect of mildronate might be partially mediated by CrAT inhibition and following regulation of cellular energy metabolism pathways.

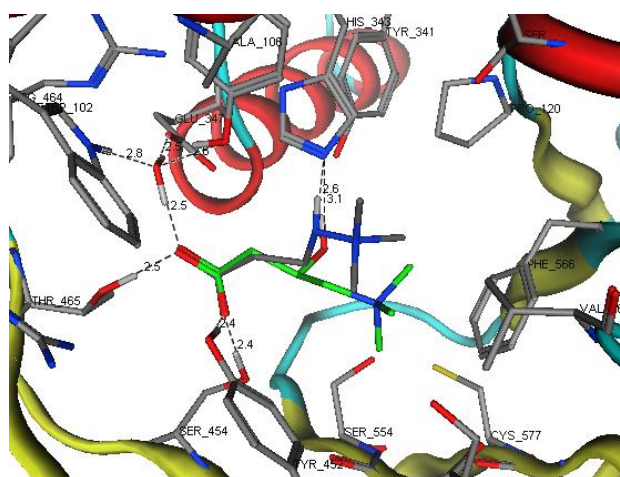


Fig. 12. Carnitine positioning in the crystal structure of carnitine-CrAT complex and comparison with the docked conformation of mildronate. Carnitine carbons are shown in green and mildronate carbons are shown in white. All hydrogens, except those involved in hydrogen bonds, are omitted.

CONCLUSIONS

The three-dimensional structures of the proteins DnaI-N, NP_247299.1 and Nsp7 in solution have been determined *de novo* using modern NMR techniques in combination with ^{15}N or double ^{15}N and ^{13}C isotope labeling, and the solution structure of the R3H domain has been re-determined for its complex with dGMP. The chemical shifts and coordinates of the structures have been deposited in the publicly available BioMagResBank and the Protein Data Bank.

DnaI:

The structure of DnaI-N comprises four α helices and two β strands as well as a short 3_{10} helix and presents a novel fold. DnaI contains a zinc ion that probably plays a structural role. The presence of conserved hydrophobic residues in the unstructured N-terminal tail of the protein suggests a functional role for these flexible and highly solvent accessible residues, presumably in interactions with the helicase.

NP_247299.1:

The application of an automated protein structure determination protocol is demonstrated yielding a significant reduction in the spectroscopist and instrument time needed for data collection, analysis and refinement. A new protocol for X-ray and NMR structure comparison has been developed, which enabled the interpretation of local differences in the highly similar NMR and crystal structures of NP_247299.1. These differences were found in segments of higher B values in the crystal and correlate with dynamic structural polymorphisms in solution.

Nsp7:

The structure of Nsp7 contains 4 helices, helices $\alpha 2$ and $\alpha 3$ form a conserved core with their lengths, positions, and relative orientations being largely preserved in different environments, while helices $\alpha 1$ and $\alpha 4$ are adopting quite different lengths, positions in the sequence, and relative orientations. The residues 70–82 exist in a dynamic equilibrium between α -helix ($\alpha 4$) and unstructured conformations suggesting that Nsp8 could bind to a transient form of Nsp7, in which helix $\alpha 4$ is unfolded. Binding to Nsp8 would trigger the formation of the short helix $\alpha 4$ immediately after $\alpha 3$, as well as the formation of the long helix $\alpha 1$, which then takes up the position occupied by $\alpha 4$ in the pH 6.5 solution structure. In the complex, the locations of $\alpha 1$ and $\alpha 4$ are stabilized by numerous contacts with hydrophobic side chains of Nsp8.

R3H:

The R3H domain binds mononucleotide mono-phosphates, and the dissociation constants of the complexes are within the micro-molar region. It has a substantial preference for binding with deoxy-ribonucleotides over ribonucleotides, especially for the purine-containing bases. The R3H domain does not bind mononucleotide di-phosphates and tri-phosphates as well as unphosphorylated mononucleotides not present in the DNA and the RNA. The R3H domain has no affinity towards 3'-phosphorylated mononucleotides too. The binding of deoxy-guanosine mono-phosphate occurs through electrostatic interactions with the positively charged sidechains of Arg47 and Arg70, π - π stacking interactions with His51 and several hydrogen bonds. His51 could be responsible for selectivity against purine-containing

nucleotides that provide better stacking interactions and the hydrophobic side chain of Leu48 could be responsible for selectivity against deoxy-ribonucleotides.

CrAT:

A model for CrAT inhibition by the antiischemic agent mildronate has been established. Mildronate inhibits CrAT in a competitive manner through binding to the carnitine binding site, not the acetyl-CoA binding site. The bound conformation of mildronate closely resembles that of carnitine except for the orientation of the trimethylammonium group, which in the mildronate molecule is exposed to the solvent. The dissociation constant of the mildronate CrAT complex is approximately 0.1 mM, and the K_i is 1.6 mM. The results suggest that the cardioprotective effect of mildronate might be partially mediated by CrAT inhibition and concomitant regulation of cellular energy metabolism pathways.

COMPLETE LIST OF AUTHOR'S PUBLICATIONS

1. Johnson M.A., Jaudzems K., Wüthrich K. NMR structure of the SARS coronavirus nonstructural protein Nsp7 in solution at pH 6.5// *Journal of Molecular Biology*. - 2010. - Vol. 402. - p. 619-628.
2. Jaudzems K., Geralt M., Serrano P., Mohanty B., Horst R., Pedrini B., Elsliger M.-A., Wilson I.A., Wüthrich K. NMR structure of the protein NP_247299.1: comparison with the crystal structure// *Acta Crystallographica Section F*. - 2010. - Vol. 66. - p. 1367-1380.
3. Mohanty B., Serrano P., Pedrini B., Jaudzems K., Geralt M., Horst R., Herrmann T., Elsliger M.-A., Wilson I.A., Wüthrich K. Comparison of NMR and crystal structures for the proteins TM1112 and TM1367// *Acta Crystallographica Section F*. - 2010. - Vol. 66. - p. 1381-1392.
4. Serrano P., Pedrini B., Geralt M., Jaudzems K., Mohanty B., Horst R., Herrmann T., Elsliger M.-A., Wilson I.A., Wüthrich K. Comparison of NMR and crystal structures highlights conformational isomerism in protein active sites// *Acta Crystallographica Section F*. - 2010. - Vol. 66. - p. 1393-1405.
5. Jaudzems K., Kuka J., Gutsaits A., Zinovjevs K., Kalvinsh I., Liepinsh E., Liepinsh E., Dambrova M. Inhibition of carnitine acetyltransferase by mildronate, a regulator of energy metabolism// *Journal of Enzyme Inhibition and Medicinal Chemistry*. - 2009. - Vol. 24. - p. 1269-1275.
6. Jaudzems K., Zhulyenkov D., Liepinsh E. R3H domain complexes with mononucleotides// *Scientific Journal of RTU. Material Science and Applied Chemistry*. - 2009. - Vol. 20. - p. 58-63.
7. Loscha K.V., Jaudzems K., Ioannou C., Su X.-C., Hill F.R., Otting G., Dixon N.E., Liepinsh E. A novel zinc-binding fold in the helicase interaction domain of the *Bacillus subtilis* DnaI helicase loader// *Nucleic Acids Research*. - 2009. - Vol. 37. - p. 2395-2404.
8. Zinovjevs K., Jaudzems K., Liepinsh E. Determination of the 3D structure of DNA hexamers (ATGCAT, ATCGAT, TAGCTA, TACGTA) by 2D NMR spectroscopy// *Scientific Journal of RTU. Material Science and Applied Chemistry*. - 2008. - Vol. 16. - p. 99-106.
9. Jaudzems K., Liepinsh E., Bizdena E. Determination of the configuration of chiral centres in the derivatives of monosaccharides by NMR// *Scientific Journal of RTU. Material Science and Applied Chemistry*. - 2006. - Vol. 13. - p. 71-77.

ACKNOWLEDGEMENTS

This work would not be possible without the contributions of many others. I would like to thank these people for their selfless work, collaboration and sharing of knowledge.

First, I would like to thank my thesis supervisor prof. Edvards Liepinsh for his excellent guidance in my way to becoming an independent researcher with world-class experience. Only thanks to his determined efforts and influence, I had the opportunity to spend two years in the Nobel laureate's, prof. Kurt Wüthrich's Laboratory in the United States. I would also like to thank him for his trustfulness by assigning me ever more important tasks that promoted my interest and enthusiasm.

Second, I would like to thank my supervisor at The Scripps Research Institute (TSRI), prof. Kurt Wüthrich for his first-class guidance, stimulating discussions and attention that I received. It was an honour to work under the supervision of such great person.

Many personal thanks to my colleagues at TSRI Michael Geralt, Pedro Serrano, Reto Horst, Biswaranjan Mohanty, Pawel Stanczak, Bill Pedrini, Maggie Johnson, Lukas Susac, Atia-tul Wahab, Amarnath Chatterjee and Sharron Thomas for their assistance, support and enjoyable discussions, thus creating an excellent environment for dedicated work. Many of them became my close friends.

Special thanks to all my colleagues at the Latvian Institute of Organic Sythesis Dr. Osvalds Pugovičs, Dace Kārkle, Dr. Aigars Jirgensons and many others for their cooperation and support for my studies in the U.S. While studying in the U.S., I always knew that I am awaited back. Thanks to them for that.

I would like to thank all the people with whom we collaborated in preparing the publications that are included in the thesis - the groups of prof. Gottfried Otting, prof. Nick Dixon, prof. Kurt Wüthrich, Dr. Maija Dambrova and Dmitry Zhulyenkov.

I would like to acknowledge prof. Ērika Bizdēna for her contributions to my initial interest and involvement in research activities, which began with a research project on DNA fragment synthesis during my last-year studies at the Āgenskalns State Gymnasium and continued with undergraduate thesis in sugar chemistry, both under her supervision. After undergraduate studies she recommended me to the Laboratory of Physical-Organic Chemistry at LIOS. This turned out to be a very important turning point in my life.

I thank the Riga Technical University, the Latvian Institute of Organic Synthesis, the State Research Programme and the European Social Fund for financial support during my PhD studies.

Precious thanks to my mom for believing in my efforts, her loving care and support.

Finally, I thank my beloved Juta for her invaluable patience, understanding and support throughout my PhD studies and especially during my two-year long absence.

LITERATURE

1. Schaeffer P.M., Headlam M.J., Dixon N.E. Protein-protein interactions in the eubacterial replisome// *IUBMB Life*. - 2005. - Vol. 57. - p. 5-12.
2. Bruand C., Farache M., MCGovern S., Ehrlich S.D., Polard P. DnaB, DnaD and DnaI proteins are components of the *Bacillus subtilis* replication restart primosome// *Mol. Microbiol.* - 2001. - Vol. 42. - p. 245-255.
3. Velten M., MCGovern S., Marsin S., Ehrlich S.D., Noirot P., Polard P. A two-protein strategy for the functional loading of a cellular replicative DNA helicase// *Mol. Cell*. - 2003. - Vol. 11. - p. 1009-1020.
4. Ioannou C., Schaeffer P.M., Dixon N.E., Soultanas P. Helicase binding to DnaI exposes a cryptic DNA-binding site during helicase loading in *Bacillus subtilis*// *Nucleic Acids Res.* - 2006. - Vol. 34. - p. 5247-5258.
5. Chen Y.L., Park S., Thornburg R.W., Tabatabai L.B., Kintanar A. Structural characterization of the active site of *Brucella abortus* Cu-Zn superoxide dismutase: a ¹⁵N and ¹H NMR investigation// *Biochemistry*. - 1995. - Vol. 34. - p. 12265-12275.
6. Holm L., Sander C. Mapping the protein universe// *Science*. - 1996. - Vol. 273. - p. 595-603.
7. Hiller S., Fiorito F., Wüthrich K., Wider G. Automated projection spectroscopy (APSY)// *Proc. Natl. Acad. Sci. USA*. - 2005. - Vol. 102. - p. 10876-10881.
8. Hiller S., Wider G., Wüthrich K. APSY-NMR with proteins: practical aspects and backbone assignment// *J. Biomol. NMR*. - 2008. - Vol. 42. - p. 179-195.
9. Volk J., Herrmann T., Wüthrich K. Automated sequence-specific protein NMR assignment using the memetic algorithm MATCH// *J. Biomol. NMR*. - 2008. - Vol. 41. - p. 127-138.
10. Fiorito F., Herrmann T., Damberger F.F., Wüthrich K. Automated amino acid side-chain NMR assignment of proteins using (¹³C)- and (¹⁵N)-resolved 3D [(¹H), (¹H)]-NOESY// *J. Biomol. NMR*. - 2008. - Vol. 42. - p. 23-33.
11. Herrmann T., Güntert P., Wüthrich K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS// *J. Biomol. NMR*. - 2002. - Vol. 24. - p. 171-189.
12. Herrmann T., Güntert P., Wüthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA// *J. Mol. Biol.* - 2002. - Vol. 319. - p. 209-227.
13. Güntert P., Mumenthaler C., Wüthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA// *J. Mol. Biol.* - 1997. - Vol. 273. - p. 283-298.
14. Deming D.J., Graham R.L., Denison M.R., Baric R.S. Processing of open reading frame 1a replicase proteins nsp7 to nsp10 in murine hepatitis virus strain A59 replication// *J. Virol.* - 2007. - Vol. 81. - p. 10280-10291.
15. Peti W., Johnson M.A., Herrmann T., Neuman B.W., Buchmeier M.J., Nelson M., Joseph J., Page R., Stevens R.C., Kuhn P., Wüthrich K. Structural genomics of the severe acute respiratory syndrome coronavirus: nuclear magnetic resonance structure of the protein nsP7// *J. Virol.* - 2005. - Vol. 79. - p. 12905-12913.

16. Zhai Y., Sun F., Li X., Pang H., Xu X., Bartlam M., Rao Z. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer// *Nat. Struct. Mol. Biol.* - 2005. - Vol. 12. - p. 980-986.
17. Imbert I., Guillemot J.C., Bourhis J.M., Bussetta C., Coutard B., Egloff M.P., Ferron F., Gorbalenya A.E., Canard B. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus// *EMBO J.* - 2006. - Vol. 25. - p. 4933-4942.
18. Bai Y., Milne J.S., Mayne L., Englander S.W. Primary structure effects on peptide group hydrogen exchange// *Proteins.* - 1993. - Vol. 17. - p. 75-86.
19. Grishin N.V. The R3H motif: a domain that binds single-stranded nucleic acids// *Trends Biochem. Sci.* - 1998. - Vol. 23. - p. 329-330.
20. Liepinsh E., Leonchiks A., Sharipo A., Guignard L., Otting G. Solution structure of the R3H domain from human Smubp-2// *J. Mol. Biol.* - 2003. - Vol. 326. - p. 217-223.
21. Fielding L. NMR methods for the determination of protein-ligand dissociation constants// *Prog. NMR Spectrosc.* - 2007. - Vol. 51. - p. 219-242.
22. Ramsay R.R., Zammit V.A. Carnitine acyltransferases and their influence on CoA pools in health and disease// *Mol. Aspects Med.* - 2004. - Vol. 25. - p. 475-493.
23. Ramsay R.R., Arduini A. The carnitine acyltransferases and their role in modulating acyl-CoA pools// *Arch. Biochem. Biophys.* - 1993. - Vol. 302. - p. 307-314.
24. Liepinsh E., Vilskersts R., Loca D., Kirjanova O., Pugovichs O., Kalvinsh I., Dambrova M. Mildronate, an inhibitor of carnitine biosynthesis, induces an increase in gamma-butyrobetaine contents and cardioprotection in isolated rat heart infarction// *J. Cardiovasc. Pharmacol.* - 2006. - Vol. 48. - p. 314-319.
25. Simkhovich B.Z., Shutenko Z.V., Meirena D.V., Khagi K.B., Mezapuke R.J., Molodchina T.N., Kalvins I.J., Lukevics E. 3-(2,2,2-Trimethylhydrazinium)propionate (THP) - a novel gamma-butyrobetaine hydroxylase inhibitor with cardioprotective properties// *Biochem. Pharmacol.* - 1988. - Vol. 37. - p. 195-202.
26. Wang Y.S., Liu D., Wyss D.F. Competition STD NMR for the detection of high-affinity ligands and NMR-based screening// *Magn. Reson. Chem.* - 2004. - Vol. 42. - p. 485-489.