

AZƏRBAYCAN RESPUBLİKASI

Əlyazması hüququnda

ATOM YERDƏYİŞMƏ PARAMETRLƏRİNİN STATİSTİK ANALİZİ VƏ VALİDASIYASI

Ixtisas: 2424.01 – Riyazi biologiya, bioinformatika

Elm sahəsi: Biologiya

İddiaçı: **Rəfiqə Çingiz qızı Məsməliyeva**

Fəlsəfə doktoru elmi dərəcəsi
almaq üçün təqdim edilmiş dissertasiyanın

A V T O R E F E R A T I

Bakı – 2021

Dissertasiya işi AMEA-nın Molekulyar Biologiya və Biotexnologiyalar İnstitutunun Hesablama struktur biologiyası laboratoriyasında yerinə yetirilmişdir.


Elmi rəhbər: fizika-riyaziyyat üzrə fəlsəfə doktoru,
professor **Qərib Novruz oğlu Mürşüdoğ**

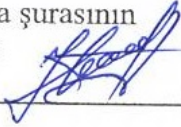
Rəsmi oponentlər: fizika-riyaziyyat elmləri doktoru, professor,
AMEA-nın müxbir üzvü
Oktay Kazım oğlu Qasımov

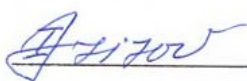
biologiya elmləri doktoru, professor
İradə Nurəddin qızı Əliyeva

biologiya elmləri doktoru, professor
Elimxan Süleyman oğlu Cəfərov

Azərbaycan Respublikasının Prezidenti yanında Ali Attestasiya Komissiyasının AMEA Molekulyar Biologiya və Biotexnologiyalar İnstitutunun nəzdində fəaliyyət göstərən BED 1.25–Birdəfəlik Dissertasiya şurası

Dissertasiya şurasının
sədri: 
biologiya elmləri doktoru, professor,
AMEA-nın həqiqi üzvü
İradə Məmməd qızı Hüseynova

Dissertasiya şurasının
elmi katibi: 
biologiya üzrə fəlsəfə doktoru, dosent
Nurməmməd Şamil oğlu Mustafayev

Elmi seminarın sədri: 
biologiya elmləri doktoru, professor,
AMEA-nın müxbir üzvü
İbrahim Vahab oğlu Əzizov

GİRİŞ

Mövzunun aktuallığı və işlənmə dərəcəsi. Atom modellərinin parametrlərinə pozisiyalar (x y z), dolğunluq və izotropik/anizotropik atom yerdəyişmə parametrləri (AYP, B qiyməti, U qiyməti, temperatur faktoru, B faktor) daxildir. Bu adlandırmalar qarşılıqlı əvəz ediləndir, lakin dissertasiyada əsasən Atom Yerdəyişmə Parametrləri (AYP) terminindən istifadə ediləcəkdir. AYP-nin validasiyasına bir çox elmi nəşrlər həsr olunduğuna baxmayaraq, onların analizinə aid çox az sayda tədqiqatlar vardır^{1,2,3}. AYP çox vaxt keyfiyyət göstəricisi kimi, modelin etibarlı hissələrinin seçilməsi üçün istifadə edilir⁴. Lakin bu prosedurlarla əlaqədar ən azı iki problem vardır^{5,6}:

1. AYP kristalın zədələri, molekulun daxilində atomların nisbi mobilliyi və həmçinin atom modelləşdirilməsi zamanı parametrlənmədəki səhvlər (məsələn pozisiyalarda yanlışlıqlar) ilə əlaqədar informasiya saxlayır.
2. AYP-nin mütləq qiyməti mənasızdır. Məsələn, saflaşdırmadan əvvəl kəskinləşdirmə və ya yayma ilə AYP-nin orta qiymətini təsadüfi qaydada dəyişmək olar. Bu yalnız hesablanan və müşahidə edilən Fürye əmsallarının arasındakı fərqi miqyas

¹ Dauter, Z., Murshudov, G.N., Wilson, K.S. International tables for crystallography / edited by M.G.Rossmann, E.Arnold. – Chester: International Union of Crystallography. Vol. F. – 2006. – p. 393-402.

² Merritt, E.A. Some Beq are more equivalent than others // Acta Crystallographica A, – 2011, 67 (6), – p. 512-516.

³ Merritt, E.A. To B or not to B: a question of resolution? // Acta Crystallographica D, – 2012, 68 (4), 468-477.

⁴ Chen, V.B. MolProbity: all-atom structure validation for macromolecular crystallography / W.B.Arendall, J.J.Headd, D.A.Keedy [et al.] // Acta Crystallographica D, – 2010, 66 (1), – p. 12-21.

⁵ Refinement and validation of macromolecular structures / R.C.Masmaliyeva, G.N.Murshudov // ANAS Transactions of Institute of Molecular Biology and Biotechnologies, – 2017, v. 1, – p. 80-93.

⁶ Statistical analysis of macromolecular B values / R.C.Masmaliyeva, G.N.Murshudov // The 11th International Conference on Bioinformatics of Genome Regulation and Structure/Systems Biology, – 2018, – p. 112.

göstəricilərinə təsir edəcəkdir. Ona görə də, AYP-nin nisbi qiymətlərindən istifadə etmək daha düzgün hesab edilir.

Makromolekulyar quruluşun hansı ayırdeTMədə alındığı da AYP qiymətlərin analizində əhəmiyyət kəsb edir. Belə ki, B faktorun 10 \AA^2 qədər fərqi aşağı ayırdeTMəli quruluşlarda o qədər də əhəmiyyətli olmasa da, yüksək ayırdeTMəli quruluşlarda dramatik effekTə malikdir. B qiymətlərinin arasındakı nisbi fərq də əhəmiyyətlidir. Arasında 10 \AA^2 qədər fərq, B faktorun 10 \AA^2 və 20 \AA^2 qiymətləri üçün daha dramatik effekTə malikdir, nəinki, 100 \AA^2 və 110 \AA^2 qiymətlərində. Buradan belə bir qənaəTə gəlmək olar ki, B faktorun analizi, həmçinin, atom modelinin etibarlı porsiyalarının seçilməsi istiqamətində də istifadə edilməsi aktual məsələdir. Qeyd etmək lazımdır ki, aşırı kəskinləşdirmə saflaşdırma zamanı AYP-ə mənfi qiymətlər verməklə onları fiziki nöqteyi nəzərdən mənasız edə bilər. Digər tərəfdən isə, həddən artıq yayılma elektron sıxlığı xəritəsində detalların itirilməsinə səbəb olar. Yuxarıdakıları nəzərə alaraq, aydındır ki, B faktorun mütləq qiyməti əvəzinə, onun paylanmasında istifadə etmək lazımdır.

Bu dissertasiya işində izotropik temperatur faktorunun statistik analizi aparılmışdır. Əvvəlcə AYP-nin paylanması Sürüşən tərs qamma paylanması (STQP) kimi modelləşdirilir və STQP-nin parametrləri atom modellərinin analizinə tətbiq edilir. Daha sonra, B faktor və ayırdeTMədən asılı zirvə hündürlüyü paylanması analiz edilir. Bu model bütün ZVB-ə tətbiq olunmuşdur. AYP-nin multimodal paylanmaları STQP-lərinin qarışığı modeli kimi modelləşdirilib ayrıca analiz edilmişdir. Atomların mərkəzində zirvə hündürlüyü fərqi istifadə etməklə aparılan lokal analiz ZVB-dəki yüksək ayırdeTMəli bir sıra quruluşlara tətbiq edilərək yanlış yerləşdirilmiş və yanlış parametrlənmiş atomların axtarışı üçün istifadə edilmişdir. Bu metodika ilk dəfə tətbiq edilərək makromolekulyar quruluşların validasiyası aparılmışdır. Hal-hazırda kəskinləşdirmə parametri kimi B faktorun orta qiymətindən istifadə edilir. STQP modelini tətbiq etməklə, B faktorun orta qiymətindən deyil, statistik paylanmasından istifadə etmək mümkün olacaqdır. Lokal analizin əsasında duran nəzəri yanaşmalar Wang⁷ tərəfindən göstərilmişdir, lakin validasiya üçün

⁷ Wang, J. Determination of chemical identity and occupancy from experimental

istifadə edilməmişdir. Bundan başqa, adəçəkilən məqalədən fərqli olaraq, dissertasiya işində bu paylanmanın ayırdetmədən asılılığından effektiv istifadə edilmişdir.

Tədqiqatın məqsəd və vəzifələri. Təqdim olunan dissertasiya işinin əsas məqsədi AYP-nin paylanmasının statistik modelinin qurulması və bu modeldən istifadə edərək ZVB-dəki quruluşların validasiyasını həyata keçirmək, eləcə də, atomların nisbi dolğunluq qiymətlərinin analizini həyata keçirərək atom quruluş modelindəki lokal səhvləri aşkarlamaqdır.

Qarşıya qoyulan məqsədlərlə əlaqədar olaraq, aşağıdakı tədqiqatın yerinə yetirilməsi planlaşdırılmışdır:

- AYP-nin statistik paylanmasının modelləşdirilməsi;
- AYP-nin statistik paylanması parametrlərindən istifadə etməklə yeni validasiya metodikasının hazırlanması;
- Hazırlanmış yeni metod ilə ZVB-dəki bütün quruluşların analizi;
- Atom quruluş modelində lokal səhvlərin tapılmasına əsaslanan metodun hazırlanması və ZVB-yə tətbiq olunması;
- Hazırlanmış metodikanın proqram təminatı şəklində inkişaf etdirilib elmi cəmiyyətin istifadəsinə verilməsi.

Müdafiəyə çıxarılan əsas müddəalar:

- I. Atom yerdəyişmə parametrlərinin paylanması sürüşən tərs qamma payanmasına uyğun gəlir.
- II. Multimodal atom yerdəyişmə parametrləri paylanması sürüşən tərs qamma payanmalarının qarışığı kimi modelləşdirilə bilər.
- III. Vizual cəhətdən yaxşı alınan elektron paylanmaların hesablanması və atom modellərin qurulması üçün atom yerdəyişmə parametrlərinin mütləq qiymətinin əvəzinə paylanmasından istifadə etmək daha optimaldır.
- IV. Atom yerdəyişmə parametrlərinin paylanmasının parametrlərindən istifadə etməklə hazırlanan validasiya strategiyası makromolekulyar quruluşların depositləşdirmədən əvvəl və sonra yoxlanılmasında əhəmiyyətli validasiya metodudur.
- V. Ayırdetmədən və atom yerdəyişmə parametrlərinin qiymətlərindən

asılı olan atom yerdəyişmə parametrlərinin müqayisəli analizi atom modellərində səhvlərin tapılmasına imkan verir.

Tədqiqatın elmi yeniliyi. Dissertasiya işində ilk dəfə olaraq:

- Atom yerdəyişmə parametrlərinin ehtimal paylanması Sürüşən tərs qamma paylanmasına uyğun gəldiyi göstərilmişdir.
- Atom yerdəyişmə parametrlərinin statistik analizinə əsaslanan, makromolekulyar quruluşların atom modellərin validasiyasını həyata keçirən yeni metod hazırlanmışdır.
- Sürüşən Tərs Qamma Paylanmasına Ehtimalın maksimallaşdırılması alqoritmi tətbiq edilmişdir.
- Atomların və liqandların atom yerdəyişmə parametrlərinin ətraf atomlar ilə müqayisəsinə əsaslanan “Lokal analiz” adlandırdığımız validasiya metodu inkişaf etdirilmişdir və zülal verilənlər bankındakı zülal quruluşlarına tətbiq edilmişdir.

Tədqiqatın nəzəri və praktiki əhəmiyyəti. AYP-nin statistik paylanması qurulması həm fundamental, həm də tətbiqi əhəmiyyətə malikdir. STQP-nin parametrlənməsi effektiv kəskinləşdirmə dərəcəsinin müəyyənləşdirilməsi üçün fundamental əhəmiyyət kəsb edir. Dissertasiya işi çərçivəsində hazırlanan *ToBvalid* proqram təminatı zülal quruluşlarının ZVB-ə yerləşdirilməmişdən əvvəl müəlliflər tərəfindən və sonra istifadəçilər tərəfindən validasiya metodu kimi istifadə edilməyə yönəlmişdir.

Əprobasiyası və tətbiqi. Dissertasiya işinin əsas nəticələri bir çox beynəlxalq konfranslarda işıqlandırılmışdır: 2018-ci ildə Rusiya Federasiyasının Novosibirsk şəhərində keçirilmiş “Bioinformatics of Genome Regulation and Structure\Systems Biology” adlı 11-ci beynəlxalq multikonfransda “Proteomics” sessiyasında “Statistical analysis of macromolecular B values” mövzusunda poster təqdimatı, 2019-cu ildə Avstriyanın Vyana şəhərində keçirilmiş 32-ci Avropa kristalloqrafları görüşündə “Validation, errors and noise in macromolecular crystallography” mikrosimpoziumunda “Analysis and validation of macromolecular B values” mövzusunda poster təqdimatı, 2019-cu ildə AMEA Mikrobiologiya İnstitutunda gənclər

gününə həsr edilmiş “Müasir Biologiyanın Aktual Problemləri” mövzusunda elmi-praktiki konfransda “Statistical analysis of macromolecular B values” mövzusunda şifahi çıxışla və 2020-ci ildə Böyük Britaniyanın Nottingem şəhərində keçirilən “CCP4 Study weekend” adlı konfransda “Local and Global Analysis of Macromolecular B values” mövzusunda şifahi çıxışla təqdimat. Həmçinin dissertasiya işinin ayrı-ayrı hissələri AMEA Molekulyar Biologiya və Biotexnologiyalar İnstitutunun seminarlarında təqdim edilmişdir.

Tədqiqat işinin əsas məhsulu olan *ToBvalid* proqram təminatı aşağıdakı keçiddə açıq lisenziya ilə istifadəçilərin ixtiyarına verilmişdir: <https://github.com/tobvalid/tobvalid>

İşin nəşri. Dissertasiya işinə aid 7 məqalə və 4 tezis olmaqla, 11 əsər çap edilmişdir.

Dissertasiyanın quruluşu və həcmi. Dissertasiya işi giriş, 5 fəsil, yekun, nəticə və istifadə edilmiş ədəbiyyat siyahısından ibarət olmaqla ümumi həcmi 152 səhifə təşkil edir. İşdə 45 şəkil və 10 cədvəl verilmişdir. Dissertasiya işində 247 ədəbiyyat mənbəsinə istinad edilmişdir.

2. TƏDQIQATIN MATERIAL VƏ METODİKASI

Tədqiqatın materialı olaraq, ZVB-dan rentgen kristalloqrafiyası metodu ilə alınmış molekulyar quruluşlar yüklənmiş və Refmac5 proqramı ilə yenidən saflaşdırılaraq tədqiqatlara hazırlanmışdır. ZVB-dəki quruluşların müxtəlif proqramlarla saflaşdırıldığını nəzərə alaraq, eyni bir proqram təminatı ilə yenidən saflaşdırılma analizlərin obyektivliyinin artırılması üçün faydalıdır. Dissertasiya işində 2019-cu ilin Noyabrına qədər yerləşdirilmiş, Rentgen kristalloqrafiyası metodu ilə öyrənilmiş 127,708 molekulyar quruluş lokal kompüterlərə yüklənmişdir. Növbəti analizlər üçün ayırdetməsi 1.5 və 3 Å arasında olan molekulyar quruluşlar seçilib ayrılmışdır. Ayırdetmənin bu intervalının seçilməsinin iki səbəbi var: 1) irəli sürülən hipotezi yoxlamaq üçün etibarlı quruluşların istifadə edilməsi; 2) izotropik AYP-dən istifadə edilməsi (1.5 Å və daha yüksək ayırdetmədə AYP-də anizotropiya müşahidə olunur). Qeyri-kristalloqrafik simmetriya

ilə əlaqədar məhdudlaşdırmaların nəticələrin obyektivliyinə təsir etməsinin qarşısını almaq üçün virus və viral zülallar analizlərdən kənarlaşdırılmışdır və qalan molekulyar quruluşlar CCP4 proqramlar paketinə daxil olan *Refmac5* proqramı ilə məhdudiyətli saflaşdırma rejimində yenidən saflaşdırılmışdır. Bunlardan 90840 molekulyar quruluşların avtomatik yenidən saflaşdırılması mümkün olmuşdur. Bunlardan 13903 quruluşun AYP paylanması multimodalıq müşahidə edildiyi üçün ayrıca analiz edilmişdir (Cədvəl 1).

Cədvəl 1.
Nümunə üçün istifadə olunan molekulyar quruluşların xülasəsi

pdb id kodu	Nümunə	Ayır- etmə (Å)	<i>R</i> y.s. əvvəlki	<i>R</i> y.s. sonrakı	<i>R</i> _{free} y.s. əvvəlki	<i>R</i> _{free} y.s. sonrakı	α	β	B_0 (Å ²)
1FCE	STQP	2	0.165	0.129	0.210	0.164	3.60	32.00	6.27
3AD4	Aşırı- kəskinləşdirmə	2.2	0.169	0.176	0.256	0.252	2.54	13.52	-2.20
2BP7	Kənar çıxma hal	2.9	0.227	0.193	0.254	0.245	7.83	308.21	-3.58
2GRM	Fərqli sayda rabitələr dəstləri (bimodalıq)	2.8	0.256	0.219	0.287	0.274	2.84	178.92	13.00
1NJK*	Zədəli domen (bimodalıq)	1.9	0.214	0.189	0.240	0.221	1.59	20.35	8.46
6ET7	Bimodal hala misal	2.85	0.216	0.230	0.275	0.296	4.52	299.7	0.24
4RQZ	Bimodal hala misal	2.40	0.196	0.202	0.226	0.239	2.63	110.94	26
5TU8	Üç moda olan hala misal	2.33	0.210	0.214	0.246	0.255	3.49	120.58	2.31
2WXU	Yüngül atomlar. Ağır atomlar	1.80	0.175	0.175	0.216	0.216	4.69	95.37	3.58
2ZBL	Ağır atomlar	1.6	0.152	0.135	0.182	0.160	4.85	51.89	1.32
3D3I	Ağır atomlar	1.78	0.169	0.141	0.201	0.176	2.79	23.09	3.81
1ZNZ	Zirvə hündürlüyü paylanması ağır atom	2.5	0.179	0.178	0.257	0.257	4.64	145.21	0.68
6D0J	Aşırı yayılma	3	0.24	0.24	0.26	0.292	4.79	297.52	87.54
4MOA	Modeldə səhvlər	2	0.216	0.19	0.237	0.214	1.85	17.17	9.8
5OBB	Böyük alfa	2.65	0.168	0.147	0.174	0.179	12.24	505.79	8.68
5X1O	Liqand	1.9	0.225	0.255	0.276	0.3	3.46	60.54	3.69
5ORJ	Liqand	1.99	0.2	0.209	0.221	0.244	3.96	149.25	14.01
6B9B	Liqand	1.8	0.135	0.14	0.162	0.166	3.69	42.43	10.37

*Cədvəldə bimodal hallara aid nümunələrdə unimodal hal üçün parametrlənmə verilmişdir. Dissertasiyanın "Nəticələr" hissəsində multimodal hallar üçün parametrlənmə təsvir ediləcəkdir.

Lokal analizin metodikası

Zirvə hündürlüyü paylanması. Kristalloqrafik və krio-EM təcrübələrində model qurulması üçün istifadə edilən 3D xəritələr uyğun olaraq elektron və elektrostatik potensialın sıxlığına uyğun gəldiyi üçün, verilmiş ayırdetmədə, B faktorun qiymətinin atomun mərkəzində zirvə hündürlüyünün qiymətinə necə təsir etməsinin öyrənilməsi əhəmiyyət kəsb edir. Zirvə hündürlüyü atomun tipindən, dolğunluqdan, ayırdetmədən və B faktordan asılıdır. Müxtəlif atomların zirvə hündürlüyünü müqayisə etmək üçün atom tiplərini və dolğunluğu dissertasiya işində nəzərə alınmamışdır^{8,9}. Bütün atomlar nöqtəvi atom kimi nəzərdən keçirilir. Aşağı ayırdetmədə və atomun mobilliyinə cavabdeh olan B faktorun yüksək qiymətlərində sıxlıq yayılmış olur. Bu, sıxlıq xəritəsinin atomun mərkəzindəki qiymətlərinə təsir etmiş olur. Nöqtəvi atomun mərkəzindəki sıxlıq aşağıdakı kimi hesablanabilir:

$$\rho(0) = 4\pi \int_0^{s_{max}} |s|^2 e^{-\frac{B_{mod}|s|^2}{4}} ds = \frac{8\pi}{B_{mod}} \left(-s_{max} e^{-\frac{B|s_{max}|^2}{4}} + \sqrt{\frac{\pi}{B_{mod}}} \operatorname{erf} \left(\frac{\sqrt{B_{mod}} s_{max}}{2} \right) \right)$$

Formuldan görüldüyü kimi, atomların mərkəzindəki sıxlıq ayırdetmədən və B faktordan asılıdır. Zirvə hündürlüyü paylanmasını qurmaq üçün zirvə hündürlüyünün qiymətlərini normallaşdırırıq, yəni, hər qiyməti maksimum qiymətə bölürük. Buna görə də bütün qiymətlər (0, 1] parçası üzrədir.

Lokal analiz. Əgər qonşu atomların B qiymətləri bir-birindən kəskin fərqlənirsə, bunun bir neçə səbəbi ola bilər:

1. atom (və ya atomlar) modeldə düzgün yerləşdirilməmişdir;
 2. atom (və ya atomlar) modeldə düzgün parametrlənməmişdir.
- Məsələn, doğunluq və/və ya atom tipi yanlış mənimsədilmiş ola bilər;

⁸ Resolution and resolvability in one, two and three dimensions / G.M.Gasimova, R.C.Masmaliyeva, G.N.Murshudov // Journal of Life Sciences and Biomedicine, - 2019, - 1(74), No 1, - p. 29-36.

⁹ Resolution and resolvability in macromolecular crystallography / G.M.Gasimova, R.C.Masmaliyeva, G.N.Murshudov // International Conference Dedicated to the 90th Anniversary of Academician Azad Mirzajanzade, - 2018, - p. 586-587.

Hər bir atomun qonşularının tapılması üçün hücrə-əlaqəli siyahılar metodunun implementasiya edildiyi GEMMI kitabxanasından istifadə edilmişdir. Əgər bir atomun (liqandın) üç və daha artıq qonşusu varsa, bu atom (liqand) üçün öz qonşularına nəzərən nisbi dolğunluq hesablanır. Nisbi dolğunluq ümumi sıxlıq formulu ilə hesablanır¹⁰:

$$c = \left(\frac{B_2}{B_1}\right)^{\frac{3}{2}} \frac{-s_{max}\sqrt{B_1}e^{-\frac{B_1 s_{max}^2}{4}} + \sqrt{\pi} \operatorname{erf}\left(\frac{\sqrt{B_1} s_{max}}{2}\right)}{-s_{max}\sqrt{B_2}e^{-\frac{B_2 s_{max}^2}{2}} + \sqrt{\pi} \operatorname{erf}\left(\frac{\sqrt{B_2} s_{max}}{2}\right)}$$

Yuxarıdakı düstur hər bir atom (c_0) və onun qonşu atomları (c_1 , c_3) üçün hesablanır. Tutaq ki, c_0 , c_1 və c_3 uyğun olaraq qonşu atomların AYP-nin medianı, birinci və üçüncü kvartilləridir. Əgər $c_0 > 1.2$ və $c_1 > 1.01$ olarsa, atom potensial ağır, $c_0 < 0.8$ (liqandlar üçün 0.7) və $c_3 < 0.99$ olarsa, potensial yüngül hesab edilir. Atomlar və liqandlar üçün local analizin alqoritmi eynidir. Fərq ondadır ki, liqandlar 1-dən çox atomdan ibarət olduğu üçün, liqanda bütöv bir vahid kimi baxılır və liqandın B faktoru kimi onun atomlarının B faktorlarının medianı istifadə edilir.

Qlobal analizin metodikası

STQP-nin ilkin qiymətləri momentlər üsulu ilə, yekun qiymətləri isə Fişer qiymətləndirmə ilə maksimal oxşarlıq metodu vasitəsilə hesablanmışdır. Multimodal hallarda modaların sayının tapılması üçün Python scipy kitabxanasından istifadə etməklə Silverman metodu tətbiq edilmişdir. Əvvəlcə zirvə hündürlüyü paylanmasında modaların sayı tapılır və bu paylanmaya Qaus qarışıq modeli tətbiq edilir, daha sonra buradan alınan nəticədən ilkin verilənlər kimi istifadə edilərək STQP-ə gözləntinin maksimallaşdırılması alqoritmi tətbiq edilərək multimodal AYP paylanması üçün parametrlər hesablanır^{11,12,13}.

¹⁰ Local and global analysis of macromolecular Atomic Displacement Parameters / R.C.Masmaliyeva, K.H.Babai, G.N.Murshudov // Acta Crystallographica Section D: Structural Biology, – 2020, 76(10), – p. 926-937.

¹¹ Analysis and validation of B values of macromolecular structures / R.C.Masmaliyeva, G.N.Murshudov // Acta Crystallographica A-foundation and advances, – 2019, 75(5), – p. e166.

NƏTİCƏLƏR VƏ ONLARIN MÜZAKİRƏSİ

3. Lokal analizin nəticələri

Əgər qəbul etsək ki, vahid hücrə boyu küyün səviyyəsi az-çox sabitdir, iddia edilə bilər ki, lokal siqnal-küy nisbəti elektron sıxlığının hündürlüyünün orta qiymətindən asılıdır və bu sonuncu da öz növbəsində molekulun lokal mobilliyindən asılıdır. Buna görə də aşağıdakı hallar mümkündür:

1) Atomlar yanlış pozisiyada yerləşdirilsə, saflaşdırma zamanı elektron sıxlığının yoxluğuna və ya çox zəif olmasına cavab olaraq, B qiymətləri kəskin şəkildə artır - beləliklə, bu hissələrdə siqnal-küy nisbəti sıfıra yaxınlaşır və ya sıfıra bərabər olur.

2) Əgər iki və daha artıq domentlə/subvahidlər müxtəlif sayda molekul daxili və/və ya kristal daxili əlaqələrə malikdirsə, onlar müxtəlif AYP-yə və beləliklə də, müxtəlif mobillik səviyyəsinə malik olacaqlar. Bu da öz növbəsində siqnalın küyə nisbətini azaldaraq belə regionların interpretasiyasını cətinləşdirir.

Hər iki halda AYP-nin paylanması multimodallıq müşahidə ediləcək və siqnal-küy nisbətləri fərqli olacaq. Bu o deməkdir ki, bəzi kristal quruluşlar üçün lokal siqnal-küy nisbəti və buna görə də lokal ayırdetmə vahid hücrə boyu dəyişir və atomların B faktorunun paylanması eyni moda altında birləşmiş klasteri üçün sabitlik nümayiş etdirir.

Dissertasiya işi çərçivəsində ayırdetməsi 2 Å və daha yuxarı olan 2000-ə yaxın molekulyar quruluş üzərində lokal analiz aparılmışdır və 628 quruluşda potensial ağır atomlar tapılmışdır və onların elektron sıxlığı xəritələri dərinədən analiz edilmişdir.

¹² Analysis and validation of macromolecular B values / R.C.Masmaliyeva, G.N.Murshudov // Acta Crystallographica Section D: Structural Biology, -2019, -75(5), – p. 505-518.

¹³ Statistical analysis of macromolecular B values / R.C.Masmaliyeva, G.N.Murshudov // Gənclər gününə həsr olunmuş “Müasir Biologiyanın Aktual Problemləri” mövzusunda elmi-praktiki konfrans, – Bakı: – 2019, – p. 48-49.

Potensial yüngül atomların görünmə dərəcəsi aşağıdır və ya onlar radiasiya zədəsinə məruz qalmışdılar. Atom öz qonşu atomlarına nisbətən daha yüksək B qiymətinə malikdirsə, alqoritm onu potensial yüngül atom kimi qeydə alır. Dissertasiya işində bir nümunə üzərində göstərilmişdir ki, potensial yüngül atom kimi qeydə alınan atomun daxil olduğu amin turşu qalığında rotamer yanlış modelləşdirilmişdir və burada yenidən-qurmaya ehtiyac duyulur. Yenidən-qurmadan sonra alqoritm bu atomu kənaraçıxma kimi qeydə almamışdır¹⁴.

Yanlış modelləşdirilmiş potensial ağır atomlar yanlış parametrlənmə nəticəsində meydana çıxır. Bir sıra quruluşlarda yanlışlıqla su molekulu kimi modelləşdirilmiş ağır atomlar müəyyən edilmişdir. Bəzən metal atomları su kimi modelləşdirilə bilirlər. Dissertasiyada potensial ağır atom kimi qeydə alınmış bir sıra yanlış modelləşdirilmiş metal atomları və xlorə aid misallar verilmişdir.

Lokal analiz həmçinin liqandlara tətbiq edilmişdir. Bu yanaşmada bütün liqand bir vahid kimi nəzərə alınır. Alqoritm aşağıdakı kimidir:

- I. Liqand atomlarının B qiymətlərinin medianı hesablanır;
- II. Liqandın hər bir atomunun qonşu atomlarının siyahısı qurulur və bu siyahıya daxil olan atomların B qiymətlərinin medianı hesablanır. Burada bir neçə məqam nəzərə alınmalıdır. Liqandın ətraf mühitinin siyahısı hazırlanarkən, bu liqandın öz atomları və ətraf mühitdən təkrarən siyahıya salınan atomlar siyahıdan kənarlaşdırılmalıdır. Belə ki, liqandın atomlarının qonşu atomların həm də liqandın öz atomları daxil ola bilər və yaxud da, liqandın müxtəlif atomlarının eyni qonşu atomları ola bilər, lakin siyahıda təkrarlar olmamalıdır;
- III. Yuxarıda hesablanan median qiymətlərindən istifadə edərək, lokal analizə daxil olan ümumi sıxlıq formulu ilə liqand üçün optimal dolğunluq qiyməti hesablanır;
- IV. Yuxarıda hesablanan median qiymətləri həm də zirvə

¹⁴ Outlier detection in atomic temperature factor - B value distribution / R.C.Masmaliyeva, G.N.Murshudov // PANAS Biological and Medicine Sciences, -2018, -73(2), – p. 25-32.

hündürlüyü formuluna tətbiq edilir;

- V. Proqram təminatı çıxış verilənləri olaraq, hesablanan optimal dolğunluq, zirvə hündürlüyü qiyməti, eləcə də, liqandın və ətraf mühitin B qiymətləri (medianları) çap edilir.

Dissertasiya işi çərçivəsində analiz edilmiş bütün molekulyar quruluşların liqandları analiz edilmişdir. Bunlarda bir neçəsi nümunə olaraq dissertasiya işində verilmişdir. Analizlərdə məlum olmuşdur ki, bəzi liqandlar əslində kristalda mövcud deyil və yaxud da modeldə verildiyindən daha aşağı dolğunluq ilə təmsil olunmuşdur. Qeyd etmək lazımdır ki, dissertasiya layihəsinin məqsədi hər hansı quruluşları tənqid etmək deyil, hazırlanmış metodun əhəmiyyətini nümayiş etdirməkdir.

Dissertasiya işində aşkar edilmiş potensial yüngül/ağır atomlara aid nümunə kimi aşağıdakılar verilmişdir: 2WXU pdb-də A zənciri, 131-ci amin turşusu (ILE) CD1 atomu 0.64 optimal dolğunluq ilə potensial yüngül atomdur. Səbəb odur ki, bu amin turşu qalığı yanlış rotamer ilə modelləşdirilmişdir.

Hazırlanmış metod vasitəsilə həmçinin yanlış modelləşdirilmiş potensial ağır atomlar da müəyyən edilir. Yanlış modelləşdirilmiş potensial ağır atomlar yanlış parametrlənmə nəticəsində meydana çıxır. Bir sıra quruluşlarda yanlışlıqla su molekulu kimi modelləşdirilmiş ağır atomlar müəyyən edilmişdir. Bəzən metal atomları su kimi modelləşdirilə bilirlər. Dissertasiyada 2ZBL id koda malik olan molekulyar quruluş modelində su kimi modelləşdirilmiş metal atomuna misal göstərilmişdir. Burada 515F (F zəncirində 515-ci qalıq) HOH molekulu kimi modelləşdirilmiş potensial ağır atomdur. Bu atom 6 oksigen atomu ilə rəbitələnmiş ideal oktahedrdir. Atomun B faktorunun qiyməti 7 \AA^2 , ətraf atomların B faktorlarının qiymətlərinin medianı 15 \AA^2 -dir və optimal dolğunluq 1.37 kimi hesablanmışdır. Kristallaşdırma şəraiti haqqında izahda qeyd edilir ki, bufferdə MgCl_2 istifadə edilmişdir. Bu pozisiyada Mg^{2+} olması daha çox həqiqətə uyğundur. Bu atom ilə ətraf atomların analiz göstərmişdir ki, məsafələr 2.09-2.2 Å arasındadır. Mg^{2+} və O arasında ideal məsafə 2.06 Å-dir. Bunu nəzərə alaraq, nəticə çıxarmaq olar ki, Mg^{2+} bu pozisiyaya uyğun gəlir. Bu pozisiyada yenidən qurma ilə Mg yerləşdirib bir neçə dövr yenidən-

saflaşdırma tətbiq etdikdən sonra atomun B faktorunun qiyməti $13 A^2$, onun üçün hesablanmış optimal dolğunluq isə 1.09 olmuşdur.

Modeldə buraxılmış xlor atomlarının düzgün yerləşdirilməsi vacibdir. Təəssüf ki, xlor atomlarının bu yolla dəqiq yerləşdirilməsi hələ mümkün deyil və yalnız intuitiv xarakter daşıyır. Bir çox hallarda bu atomları modeldə səhvən buraxılır. Onların müəyyən edilməsi üçün *ToBvalid* proqramına daxil olan lokal analizdən və müəyyən etdikdən sonra əlavə edilməsi üçün Coot proqramından istifadə etmək olar. Elektron sıxlığından bu atomun xlor olub-olmamasını müəyyən etmək çətin olsa da, kristallaşdırma şərtlərindən ağır atomun yerinə Cl (xlor) atomunun yerləşdirilə biləcəyi haqqında məlumat verə bilər. ZVB-də 3D3I id koda olan molekulyar quruluş modeli buna misal ola bilər. Bu molekulun A zəncirin 2003 su molekulu optimal dolğunluğu 1.46 olan potensial ağır atomdur. Kristallaşdırma şərtlərində qeyd edilir ki, kristala maqnesium-xlorid əlavə edilmişdir. Buradan intuitive də olsa belə nəticəyə gəlmək olar ki, bu atom Cl (xlor) atomu ola bilər. Hal-hazırda bu cür yoxlamalar nəticəsində müşahidə edilən bu ağır atomların xlor olması yalnız bioloqun biliklərinə əsaslanıb intuitiv xarakter daşısa da, gələcəkdə bu halları dəqiq təsbit edə bilən metodlar olacaqdır. Dissertasiya işi çərçivəsində hazırlanmış metodun məqsədi isə belə kənarçıxmaları tapıb işarələməkdir.

Metal liqandlara malik molekulyar quruluşlarda metal atomu düzgün modelləşdirilibsə, belə atomlar *ToBvalid* proqramının lokal analizi tərəfindən potensial ağır atom kimi qeydə alınmır. Məhz ona görə ki, bu atomlar düzgün modelləşdirilmiş və parametrlənmişdir. Analiz edilən quruluşlar arasındakı quruluşlardan birində metal atomu ağır atom kimi qeydə alınmışdır. ZVB-də 2WXU id koda malik olan quruluş modelində A zəncirinin 1377 Ca atomu isə 1.37 optimal dolğunluqla potensial ağır atomdur. Kristallaşdırma şərtlərindən görünür ki, $CdSO_4$ (kadmium-sulfat) əlavə edilmişdir. Burada iki ehtimal ola bilər: 1) bu pozisiyada Cd (kadmium) olmalıdır lakin bu atom Ca (kalsium) kimi modelləşdirilib; 2) Burada su ilə yarım-dolğunluqlu Cd (kadmium) yanlış olaraq Ca (kalsium) kimi modelləşdirilmişdir. Kalsiumun 20 elektronu vardır və 1.38 optimal dolğunluqla ($20 \cdot 1.37 = 27.6$) bu atomun yerində olan

atomun təxminən 28 elektornu olmalıdır. Kadmiumun 48 elektronu var. Su ilə yarım-dolğunluqlu Cd (kadmium) 28 elektron verir:

$$48 * 0.5 = 24 \text{ (Cd 0.5 dolğunluq)}$$

$$24 + 4 = 28 \text{ (HOH 0.5 dolğunluq } 8 * 0.5 = 4)$$

Beləliklə, qonşu atomları nəzərə almaqla, hər bir atomun və onun qonşu atomlarının mərkəzindəki zirvə hündürlüyünün müqayisəli analizi yanlış yerləşdirilmiş və yanlış parametrlənmiş (atomun tipi, dolğunluq) atomların axtarışında güclü alət kimi istifadə edilə bilər.

Dissertasiya işi çərçivəsində ümumilikdə 14176 quruluş modelində liqand analizi aparılmışdır. Bunlarda bir neçəsi nümunə olaraq dissertasiya işində verilmişdir. Burada qeyd etmək lazımdır ki, dissertasiya işinin məqsədi bütün zülal molekullarının analizi deyil, bu quruluşların validasiyasını həyata keçirməyə imkan verən metodikanın və proqram təminatının işlənilib hazırlanmasıdır. Analiz nəticəsində, bir sıra hallarda liqandların natamam dolğunluğa malik olduqları məlum olmuşdur. Dissertasiya işində aşkar edilmiş potensial yüngül liqandlara aid nümunə kimi Cədvəl 2-də verilmişdir.

Cədvəl 2

Liqand validasiyasında nümunə kimi istifadə edilən uruluşların liqand analizinin nəticələri

PDB kodu	5X1O	5ORJ	6B9B
Ayırdetmə (Å)	1.9	1.99	1.8
Liqand, nömrəsi, zəncir	I3P, 201, A	I6P, 407, A	MAL, 807, B
Optimal dolğunluq (ümumi sıxlıq üsulu)	0.12	0.21	0.41
Optimal dolğunluq (zirvə hündürlüyü)	0.09	0.11	0.27
Liqandın B qiymətlərinin medianı (Å ²)	125	252	99
Ətraf mühitin B qiymətlərinin medianı (Å ²)	12	53	37

Bu nümunələr göstərir ki, liqandların onların ətraf atomları ilə müqayisəli AYP analizi validasiyada əhəmiyyətli rola malik ola bilər və yanlış və ya zədəli liqandların identifikasiyasında faydalıdır.

4. Qlobal analizin nəticələri

Qlobal analizdə ZVB-də olan zülalların AYP-nin paylanması

STQP və STQP-lərin qarışığı kimi modelləşdirilmişdir¹⁵. Bu statistik paylanmaların və onların parametrlənməsinin əsasında validasiya metodologiyası hazırlanmışdır. Bu metodologiyaya STQP və STQP-lərin qarışığı paylanmalarının parametrlərinin məqbulğunun ümumi ZVB ilə müqayisəli yoxlanılması, paylanma üçün ümumi statistik hesabatı, STQP üçün ekstrimal qiymətlərin siyahıya alınması, ayrı-ayrı zəncirlər üçün AYP-nin analizi və validasiyası daxildir. Bütün bu nəticələr *ToBvalid* proqram təminatının validasiya hesabatları şəklində istifadəçilərə təqdim edilir¹⁶.

Dissertasiya işində göstərilmişdir ki, atom modellərindəki B faktorun (AYP) paylanması Sürüşən Tərs Qamma paylanmasına uyğun gəlir. Burada aşırı kəskinləşdirmənin mənfi olduğu hallar və paylanmada çoxsaylı çox kiçik B qiymətlərinin olduğu hallar istisnadır. Birinci halda B qiymətləri yanlış hesablanmışdır, ikinci halda isə model qurulması proqramı B faktorun qiymətini müəyyənləşdirərkən bir qrup atomlara hər hansı səbəbdən təsadüfi kiçik qiyməti mənimsətmişdir. ZVB-də 1316 belə hal müşahidə edilmişdir. Bunlarda temperatur faktorun minimal qiyməti ($B_{\min} = 0.5$) əksər atomlarda təmsil olunub.

Analiz edilən bütün molekulyar quruluş modelləri üçün STQP-nin parametrləri qiymətləndirilmişdir. Validasiya metodu əvvəlcə R statistik paketində implementasiya edilmişdir, daha sonra Python proqramlaşdırma dilinə köçürülmüşdür. Müəyyən olunmuşdur ki, ZVB-dəki quruluşların əksəriyyətinin izotropik temperatur faktorunun paylanması STQP-na uyğun gəlir. ZVB-dəki yüksək ayırdetməli, etibarlı molekulyar quruluşlar üçün tipik parametrlər müəyyən edilmişdir. Bu modeli doğruldan tipik quruluşlara misal olaraq, 5ZPD id koda malik quruluş modeli göstərilmişdir: atomların sayı 10795, minimum B qiyməti 8.97, maksimum B qiyməti 173.09, orta qiymət 24.77, median 20.94, variasiya 197.25, meyillilik 3.72,

¹⁵ Validation of crystal structures of Light Harvesting Complexes of green plants / R.C.Masmaliyeva // Transactions of the Institute of Molecular Biology & Biotechnologies, ANAS, – 2020, v. 4, – p. 8-14.

¹⁶ Atom yerdəyişmə parametrlərindən istifadə etməklə Zülalların Verilənlər Bazasında olan quruluşların yoxlanılması / R.Ç.Məsməliyeva // Həyat Elmləri və Biotibb Jurnalı, 1(74), No 2, – p. 5-12.

kurtosis 22.86, 1-ci kvartil 16.52, 3-cü kvartil 28.26, alfa 3.55, beta 47.9, B_0 6.09.

Sürüşən Tərs Qamma Paylanması modelinə uyğun gəlməyən hallarda verilənlərin natamamlığı və modeldə səhvlərlə müşayiət ediləcəyi gözlənilir. Alfa parametrinin 2-dən kiçik olduğu hallar açıq-aşkar STQP deyildir. Belə hal 17 quruluş modelində müşahidə edilmişdir. Bu hallardan birinə misal 4MOA id koda malik olan quruluş modelidir: atomların sayı 5056, minimum B qiyməti 10.43, maksimum B qiyməti 237.09, orta qiymət 32.91, median 20.35, variasiya 1283.77, meyillilik 2.99, kurtosis 8.31, 1-ci kvartil 16.5, 3-cü kvartil 29.51, alfa 1.85, beta 17.17, B_0 9.8. Elektron paylanmasının analizindən (Coot proqram təminatı vasitəsilə) müəyyən edilmişdir ki, modeldə aşırı bulanıq, elektron paylanmasında zəif təmsil olunan hissə vardır ki, bu hissədəki atomların B qiymətləri yüksəkdir. Bu, modelin bu hissəsində səhvlərin olduğunu, və ya hətta bu domenin kristalda olmadığını bildirə bilər. Belə səhvlər detallı analiz edilməli və modeldə yenidənqurma aparılmalıdır. Bu hissəni modeldən silib yenidən parametrlənmə aparıldıqda, quruluş modeli Sürüşən Tərs Qamma Paylanmasına uyğun parametrlər və histogram nümayiş etdirmişdir: atomların sayı 4695, minimum B qiyməti 10.43, maksimum B qiyməti 116.83, orta qiymət: 23.8, median 19.68, variasiya: 166, meyillilik 3.15, kurtosis 13.28, 1-ci kvartil 16.25, 3-cü kvartil 26.49, alfa 3.1, beta 30.5, B_0 8.95.

Ən böyük alfa parametri 5OBB id koda malik olan quruluş modeli üçün müşahidə edilmişdir (alfa: 12.24, beta: 505.79, B_0 : 8.68). Alfa'nın böyük qiymətləri histogramın tərs qamma paylanması üçün xarakterik olmayan simmetriklilik ilə müşahidə olunur. Belə quruluşların dərinədən analizinə ehtiyac duyulur. Hazırlanmış metod bu cür quruluşların aşkarlanması üçün əhəmiyyətlidir.

B qiymətlərinin paylanması müəyyən $B=B_{\min}$ ilə sola doğru sürüşməyə malikdirsə və paylanmada çoxlu sayda atom eyni kiçik B qiymətə malikdirsə, bu o deməkdir ki, verilənlər saflaşdırmadan əvvəl aşırı-kəskinləşdirilmişdir. Bu halda α parametri kiçik olur, ya da parametrlənmə zamanı STQP-nin parametrlərinin requlyarizasiyası qeyri-stabil olur. Belə verilənlərin sayı ZVB-də çox azdır. B_0 parametrinin mənfi qiymətləri də aşırı-kəskinləşdirmədən xəbər verir.

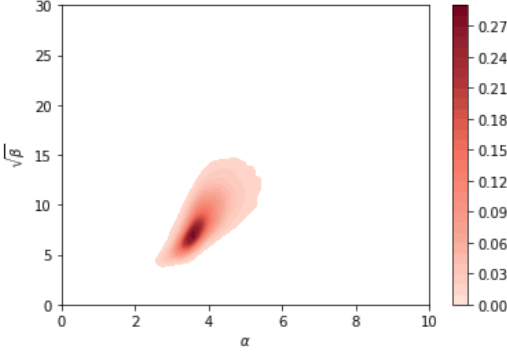
Analiz edilən quruluşlardan 2535 ədədi mənfə B₀ müşahidə edilmişdir.

Aşırı-kəskinləşdirmə halına misal olaraq 3AD4 id koda malik olan molekulyar quruluşun modelini göstərmək olar: atomların sayı 2442, minimum B qiyməti 0.5, maksimum B qiyməti 59.83, orta qiymət: 7.31, median 4.18, variasiya: 90.86, meyillilik 2.72, kurtosis 8.87, 1-ci kvartil 1.15, 3-cü kvartil 9.34, alfa 2.54, beta 13.52, B₀ -2.2. Bu misalda kiçik B qiyməti əlavə etməklə (20 Å) Refmac5 proqramı ilə yenidən saflaşdırma aparılmışdır. Kiçik yayılma əlavə etdikdən sonra STQP-nin histoqramı və parametrləri STQP-ə uyğun gəlir və alfa və beta parametrləri alfa-beta qrafikində kontura uyğun gəlir: atomların sayı 2442, minimum B qiyməti 10.01, maksimum B qiyməti 89.78, orta qiymət: 26.71, median 23.4, variasiya: 124.92, meyillilik 2.53, kurtosis 7.77, 1-ci kvartil 19.97, 3-cü kvartil 29.47, alfa 4, beta 55.67, B₀ 8.97.

Aşırı yayılma halına misal olaraq PDB kodu 6D0J olan molekulyar quruluş göstərilmişdir: atomların sayı 7611, minimum B qiyməti 96.78, maksimum B qiyməti 321.25, orta qiymət: 161.83, median 156.05, variasiya: 993.13, meyillilik 0.97, kurtosis 1, 1-ci kvartil 138.62, 3-cü kvartil 179.22, alfa 4.79, beta 297.52, B₀ 87.54. Burada STQP-də sağa doğru sürüşmə var və parametrlər alfa-beta qrafikfə konturdan kənardadır. Alfa və Beta parametrlərinin, eləcə də B₀ sürüşmə addımı parametrinin böyük qiyməti bu quruluşda “yayılma” ola biləcəyini ehtimal etməyə imkan verir. Sürüşmə addımı parametrinə uyğun olaraq kəskinləşdirmə (-60 Å) tətbiq edilərək yenidən saflaşdırma aparıldıqdan sonra parametrlər aşağıdakı kimi olmuşdur: atomların sayı 7611, minimum B qiyməti 25.8, maksimum B qiyməti 216.28, orta qiymət: 73.83, median 67.33, variasiya 856.03, meyillilik 1.18, kurtosis 1.43, 1-ci kvartil 52.09, 3-cü kvartil 88.63, alfa 4.26, beta 201.05, B₀ 14.64.

Analiz edilən bütün zülal quruluşlarının izotropik temperatur faktoru paylanması üçün alfa, beta, B₀, median, orta qiymət, minimum, maksimum, varians, meyillilik, kurtozis, 1-ci kvartil, 3-cü kvartil hesablanmışdır. Alfa və beta parametrləri zülal quruluşları üçün müəyyən qiymətlərə malik olur. Bu qanunauyğunluq zülalların quruluş modellərinin validasiyası üçün istifadə edilə bilər. Aşağıdakı şəkildə isə

alfa və beta (kvadrat kökü) parametrlərinin 2D qrafiki verilmişdir. Dissertasiya işi çərçivəsində hazırlanan *ToBvalid* proqram təminatı alfa-beta 2D kontur qrafiki generasiya edir və analiz edilən zülal quruluşunun alfa parametri və beta parametrin kvadrat kökünü həmin qrafikin üzərində yerləşdirir. Analiz edilən zülalların parametrlərin bu konturdan kənarında olarsa, bu quruluş yenidən saflaşdırma, model qurulması və yenidənqurulması üçün nəzərdə tutulmalıdır.



Şəkil. Bütün ZVB-dəki zülal quruluşlarının STQP parametrlərinin (alfa və betanın kvadrat kökü) ikiölçülü qrafiki.

ZVB-də bir sıra quruluşlar vardır ki, onların B faktorunun paylanması multimodalıq nümayiş etdirir. Analiz edilən quruluşlardan 13903 multimodal hal vardır. Bunlardan 13636 bimodaldır, 266-sı 3 modalı və biri 4 modalıdır. Multimodallığın bir neçə səbəbi ola bilər. Əsas səbəblərdən aşağıdakı kimidir:

1. Bu molekulların öz mühitində (kristal və ya kriyoEM multi-domen/multi-subvahid quruluşlar) təbii xüsusiyyəti ola bilər. Belə quruluşun müxtəlif komponentləri müxtəlif sayda rabitələr əmələ gətirir. Bu halda müxtəlif subvahidlər müxtəlif mobillik səviyyəsinə malikdir və B faktorun paylanmasında öz əksini tapır. Gözlənilir ki, hər individual struktur vahidi özünü müstəsna parametrlərə malik STQP kimi aparacaqdır. Lakin belə hallarda atomların yerləşməsinin və B qiymətlərinin saflaşdırılması qeyri-stabil olur və qeyri-dəqiq modellərin meydana çıxmasına gətirib çıxarır. Başqa cür desək, daha yüksək B qiymətlərinə malik zəncirlər ümumi paylanma üçün STQP hipotezini poza bilər.
2. Modelin bəzi hissələri (ilgək, liqand və hətta domen) yanlış yerləşdirilə bilər. Bu hissələrin strukturda yerləşməsi şübhə

altındadır. Belə hissələr molekulun funksiyası öyrənilərkən çox ehtiyatla nəzərdən keçirilməlidir.

Birinci hala misal olaraq PDB kodu 2GRM olan molekulyar quruluşu misal göstərmək olar. Bu molekulda üç zəncir vardır. C zəncirinin elektron sıxlığı A və B zəncirinin elektron sıxlığına nisbətən daha zəifdir. Bu quruluşun zəncirlərinin *Refmac5* proqramı ilə analizindən məlum olmuşdur ki, A və B zəncirləri bir-biri ilə sıx əlaqədədir, lakin C zənciri əsasən öz atomları ilə rəbitəyə malikdir. C zənciri yalnız öz atomları ilə əlaqə yaradır və assimetrik vahiddə B zənciri ilə 15 ədəd qarşılıqlı təsirə malikdir. A və B zəncirləri isə, bir-biri ilə asimetrik vahid daxilində və onlar arasında əlaqələrə malikdir. Buradan aydındır ki, C zənciri A və B zəncirlərinə nisbətən daha mütəhərrikdir və bu səbəbdən də C zəncirinin elektron paylanması daha zəifdir və uyğun olaraq da bu zəncirin atomlarının B qiymətləri daha böyükdür. Multimodallığın ikinci səbəbinə aid misal olaraq PDB kodu 1NJK olan quruluşdur. Bu halda yüksək B qiymətləri və uyğun olaraq da bulanıq sıxlığının səbəbi nasaz domendir.

ZVB-nin analizindən məlum olmuşdur ki, cəmi bir quruluş modelində AYP paylanması üç moda nümayiş etdirir. Üç modaya malik AYP paylanması nümayiş etdirən quruluşlara misal olaraq 5TU8 id koda malik olan quruluş: atomların sayı 2230, minimum B qiyməti 14.35, maksimum B qiyməti 129.68, orta qiymət 49.38, median 42.7, variasiya 725.27, meyllilik 0.93, kurtosis -0.03, 1-ci kvartil 26.89, 3-cü kvartil 64.37, alfa 3.49, beta 120.58, B_0 2.31. Bu quruluş üçün STQP-larının qarışığı modeli hazırlanmışdır. 5TU8 quruluş modelinin misalında problem ondadır ki, kristal zədəlidir və molekulun bəzi hissələrinin həddən artıq zədəli olması ucbatından kriticalın bu hissələrinin elektron sıxlığının interpretasiyası ümumiyyətlə mümkün görünür. AYP paylanmasənən ilk klasteri molekulun orta hissəsinə, digər klasterlər isə zədələnmənin başladığı hissələr olan kənarlara uyğun gəlir.

Elektron paylanması xəritəsində siqnalın səviyyəsi B qiymətlərindən asılı olduğu üçün, gözlənilir ki, müxtəlif klasterlər müxtəlif ayırdetməyə malik ola bilər. Bəzi molekulların mütəhərrik funksional domenləri atomlarının B faktorunun daha böyük qiymətlərinə malik olmaqla AYP paylanmasının multimodallığına

səbəb ola bilər. Belə hallardan biri ZVB-də 6ET7 id koda malik olan bakteriofitoxromun aktivləşmiş heterodimerinin molekulyar quruluşudur: atomların sayı 10522, minimum B qiyməti 19.15, maksimum B qiyməti 221.55, orta qiymət 81.49, median 82.21, variasiya 1171.13, meyillilik 0.3, kurtosis -0.82, 1-ci kvartil 49.02, 3-cü kvartil 108.28, alfa 4.52, beta 299.7, B_0 0.24. Bu molekulun Diguanylate cyclase (DGC) domeninin atomlarının B qiymətləri STQP-də və zirvə hündürlüyü paylanmasında ikinci piki əmələ gətirir.

4RQZ id koda malik molekulyar quruluş modelinin AYP paylanmasında iki aydın seçilən pik müşahidə edilir. Bu molekulda üç domen var. Onlardan ikisi bir-biri ilə və öz simmetriya əkiləri ilə əlaqələnilir. Bu iki domen kristalın formalaşmasında iştirak edirlər. Lakin üçüncü domen yalnız öz simmetriya əkizi ilə əlaqələr yaratmışdır. Bu domen və onun simmetriya əkilərini stabilləşdirən başqa kristal qontaktlar olmadığına görə, bu domen sərbəst hərəkət edə bilər və məhz bu səbəbdən də bu domənə daxil olan atomların B faktorlarının qiymətləri daha böyük qiymətlər alır və AYP paylanmasındakı ikinci klasteri formalaşdırır.

Beləliklə, AYP faktorun paylanması multimodalıq nümayiş etdirə bilər. Multimodal AYP paylanmasını STQP-lərin qarışığı kimi modelləşdirmək olar. Bu cür modelləşdirmə siqnal səviyyəsinin müxtəlifliyini nəzərə almaqla molekulyar quruluşların STQP hipotezi vasitəsilə validasiyasının həyata keçirilməsinə imkan verir.

5. ToBvalid validasiya proqram təminatı və kitabxanası

STQP, eləcə də, onların qarışığı modelinin parametrlənməsi verilən atom quruluş modelinin etibarlılığının qiymətləndirilməsinə, eləcə də yanlış yerləşdirilmiş və parametrlənmiş atomların müəyyən edilməsinə imkan verir. Sözügedən metodlar həm sadə skript həm də obyekt oriyentasiyalı klasslar şəklində yığılmış proqram təminatı kimi hazırlanmışdır. Bu proqram ZVB-dəki pdb faylların oxunub analiz edilməsinə imkan verir.

Hazırladığımız yeni validasiya proqramı tərəfimizdən *ToBvalid* adlandırılmışdır. Bu proqram təminatı Python proqramlaşdırma dilində yazılmışdır və Python 2.7-3.6 versiyaları

ilə uzlaşmaya malikdir. Proqram təminatı həm Unix əsaslı sistemlər olan Linux və MacOS əməliyyat sistemlərində, həm də, Windows əməliyyat sistemi mühitində işləyir. *ToBvalid* əmrlər sətri ilə işə salınır və hələ ki, heç bir qrafik intergeysə malik deyildir. Çıxış verilənləri html hesabatlar və sadə mətn faylları (txt) kimi saxlanılır. *ToBvalid* həmçinin PyPI (pip) standart paket-idarəetmə sistemi vasitəsilə dinamik kitabxana kimi istifadə edilə bilər. Bu kitabxana icra edilə bilən *tobvalid.py* proqramının bütün funksionallığına malik olub, istənilən ayrı proqramın daxili müstəqil hissəsi olaraq validasiya hesabatlarının alınmasına imkan verir. Proqram təminatının aşağıdakı kitabxanalardan asılılığı vardır: fire, numpy, scipy, gemmi, statsmodels, matplotlib, seaborn.

Proqram təminatı və kitabxana git və pip mühitlərində mövcuddur. Proqramın mənbə kodu aşağıdakı ünvanda açıq şəkildə verilmişdir: <https://github.com/ToBvalid/>. Proqramı git vasitəsilə yükləmək üçün aşağıdakı əmrdən istifadə etmək lazımdır:

```
user pip3 install git+https://github.com/ToBvalid/tobvalid
```

Pip ilə birbaşa üsulla yükləmək üçün aşağıdakı əmrdən istifadə edilir:

```
user pip install tobvalid
```

ToBvalid proqram təminatı Python3 ilə dəstəklənir və MPL-2.0 lisenziyası ilə açıq kodlarla təqdim edilir. Proqram təminatı aşağıdakı funksionallıqlara malikdir:

- AYP paylanmasının ümumi statistik analizi
- AYP paylanmasının (və ya onların qarışığının) parametrlənməsi və validasiyası
- Qonşu atomlara nəzərən potensial yüngül və ağır atomların axtarışı
- İki atomun nisbi dolğunluğunun müqayisəsi. Bu funksionallıq disseratasiya yazılan müddətdə yalnız kitabxanada mövcud olub, proqram təminatından istifadəçi interfeysinə malik deyil.

Proqramı işə salmaqdan ötrü terminalda aşağıdakı əmri vermək lazımdır:

```
tobevalid -i {Input file} -o {output directory} -m {number of modes|auto} -t {tolerance} -hr {plot resolution in dpi} -a
```

{'all'|'local'|'global'}

Əmrin qısa forması belədir:

tobvalid -i {Input file}

Sükuta görə aşağıdakı parametrlər istifadə edilir:

o – cari direktoriya

m – 1

p – heç bir (bütün lazımlı informasiya fayldadır)

Bəzən Atom Yerdəyişmə Parametrlərinin paylanması multimodal ola bilər. Bu zaman AYP paylanması STQP-lərinin qarışığı kimi modelləşdirilir və parametrlənir. Proqram təminatında sükuta görə modaların sayı 1-dir. Əgər istifadəçiyə modaların sayını müəyyənləşdirmək vacibdirsə, terminalda əmrin qısa formasına -m auto əmrini əlavə etmək lazımdır. Əgər istifadəçi yalnız lokal, qlobal və ya zəncirlərin analizini həyata keçirmək istəyirsə, əmrdə uyğun olaraq -l, -g, -c yazmaq lazımdır. Əgər bunlardan heç biri istifadə edilməsə sükuta görə bütün analizlər yerinə yetiriləcəkdir.

İnputun detalları ilə tanış olmaq üçün terminalda aşağıdakı əmri vermək lazımdır:

tobvalid -h

Geniş dokumentasiya ilə aşağıdakı ünvanla tanış olmaq olar:

https://github.com/ToBvalid/tobvalid/blob/master/doc/User_Guide.pdf

NƏTİCƏLƏR

1. Atom/liqand və ətrafı arasında ayırdetmədən və atom yerdəyişmə parametrlərindən asılı olan nisbi elektron paylanmalarının zirvə hündürlüklərinin müqayisəli analizi əsasında atomların/liqandların validasiya metodu hazırlanmışdır [3, 5, 10, 11].
2. Zülalların Verilənlər Bankındakı quruluşların atom modellərində izotopik atom yerdəyişmə parametrlərinin statistik analizi aparılmışdır və göstərilmişdir ki, izotopik atom yerdəyişmə parametrlərinin ehtimal paylanması Sürüşən tərs qamma paylanmasına uyğun gəlir [1, 2, 4, 6-9].
3. İzotropik atom yerdəyişmə parametrlərinin ehtimal paylanmasının multimodal hallarında atom yerdəyişmə

parametrlərinin ehtimal paylanması sürüşən tərs qamma paylanmalarının qarışığı modelinə uyğun gəlir. Müəyyən olunmuşdur ki, Zülalların Verilənlər Bankında rentgen kristalloqrafiyası metodu ilə alınmış 100 minə yaxın quruluşdan təxminən 10%-nin atom yerdəyişmə parametrlərinin paylanması multimodallıq nümayiş etdirir [9, 10].

4. Zülalların Verilənlər Bankındakı bütün quruluşların atom yerdəyişmə parametrlərinin sürüşən tərs qamma paylanmasının parametrlərinə əsasən hazırlanmış alfa-beta kontur qrafiki validasiya üçün proksi kimi istifadə edilir [6, 9, 11].
5. *ToBvalid* metodu atom yerdəyişmə parametrləri paylanmasının statistik analizi və validasiyasını, atomların və liqandların lokal validasiyasını həyata keçirən proqram təminatı və Python kitabxanası şəklində hazırlanmış və Mozilla Public License version 2.0 lisenziyası ilə GitHub və PyPi platformalarına yerləşdirilərək tədqiqatçıların istifadəsinə verilmişdir [10].

Dissertasiya mövzusunə aid çap edilmiş elmi nəşrlərin siyahısı

1. Refinement and validation of macromolecular structures / R.C.Masmaliyeva, G.N.Murshudov // ANAS Transactions of Institute of Molecular Biology and Biotechnologies, – 2017, v. 1, – p. 80-93.
2. Statistical analysis of macromolecular B values / R.C.Masmaliyeva, G.N.Murshudov // The 11th International Conference on Bioinformatics of Genome Regulation and Structure/Systems Biology, – 2018, – p. 112.
3. Resolution and resolvability in macromolecular crystallography / G.M.Gasimova, R.C.Masmaliyeva, G.N.Murshudov // International Conference Dedicated to the 90th Anniversary of Academician Azad Mirzajanzade, – 2018, – p. 586-587.
4. Outlier detection in atomic temperature factor - B value distribution / R.C.Masmaliyeva, G.N.Murshudov // PANAS Biological and Medicine Sciences, – 2018, 73(2), – p. 25-32.

5. Resolution and resolvability in one, two and three dimensions / G.M.Gasimova, R.C.Masmaliyeva, G.N.Murshudov // Journal of Life Sciences and Biomedicine, – 2019, 1(74), No 1, – p. 29-36.
6. Analysis and validation of macromolecular B values / R.C.Masmaliyeva, G.N.Murshudov // Acta Crystallographica Section D: Structural Biology, – 2019, 75(5), – p. 505-518.
7. Statistical analysis of macromolecular B values / R.C.Masmaliyeva, G.N.Murshudov // Gənclər gününə həsr olunmuş “Müasir Biologiyanın Aktual Problemləri” mövzusunda elmi-praktiki konfrans, – Bakı: – 2019, – p. 48-49.
8. Atom yerdəyişmə parametrlərindən istifadə etməklə Zülalların Verilənlər Bazasında olan quruluşların yoxlanılması / R.Ç.Məsməliyeva // Həyat Elmləri və Biotibb Jurnalı, – 2019, 1(74), No 2, – p. 5-12.
9. Analysis and validation of B values of macromolecular structures / R.C.Masmaliyeva, G.N.Murshudov // Acta Crystallographica A-foundation and advances, – 2019, 75(5), – p. e166.
10. Local and global analysis of macromolecular Atomic Displacement Parameters / R.C.Masmaliyeva, K.H.Babai, G.N.Murshudov // Acta Crystallographica Section D: Structural Biology, – 2020, 76(10), – p. 926-937.
11. Validation of crystal structures of Light Harvesting Complexes of green plants / R.C.Masmaliyeva // Transactions of the Institute of Molecular Biology & Biotechnologies, ANAS, – 2020, v. 4, – p. 8-14.

Dissertasiyanın müdafiəsi "29" iyun 2021-ci il tarixində saat 11⁰⁰-da AMEA-nın Molekulyar Biologiya və Biotexnologiyalar İnstitutunun nəzdində fəaliyyət göstərən BED 1.25–Birdəfəlik dissertasiya şurasının iclasında keçiriləcək.

Ünvan: AZ1073, Bakı ş., İzzət Nəbiyev küçəsi 11

Dissertasiya ilə AMEA Molekulyar Biologiya və Biotexnologiyalar İnstitutunun kitabxanasında tanış olmaq olar.

Dissertasiya və avtoreferatın elektron versiyaları AMEA Molekulyar Biologiya və Biotexnologiyalar İnstitutunun rəsmi internet saytında (<https://www.imbb.az/>) yerləşdirilmişdir.

Avtoreferat "27" may 2021-ci il tarixində zəruri ünvanlara göndərilmişdir.

Çapa imzalanıb: 21.05.2021

Kağızın formatı: A5

Həcm: 38280

Tiraj: 100