

mcgen kasutajajuhend

Heiki Kasemägi, Endel Soolo, Alvo Aabloo
and
Josh Thomas

January 11, 2004

Contents

1	<i>mcgen</i>	3
2	Disclaimer	3
3	Tööpõhimõte	3
4	Programmi struktuur	4
4.1	Programmi üldine struktuur	4
4.2	Sisendfailide lugemine	4
4.3	Genereeritava struktuuri koostamine sisendinfo põhjal	5
5	Sisend	6
5.1	<i>mcgen</i> käsurida	6
5.2	Tüübikirjeldused: <i>defs.ini</i>	7
5.3	<i>uniti</i> kirjeldus: <i>units.ini</i>	9
5.4	Ahelate struktuur: <i>chains.ini</i>	11
5.5	Jõuväli: <i>field.ini</i>	14
5.6	Üldised sisendparameetrid: <i>params.ini</i>	20
5.7	Kasutaja defineeritavad geomeetrilised piirangud: <i>userfunc.cpp</i>	23
5.8	Juhuarvude generaator: <i>randome.sed</i>	24
6	Väljund	24
6.1	Logifail: <i>mcgen.log</i>	24
6.2	XYZ: <i>mcgen.xyz</i>	25
6.3	DLPOLY CONFIG: <i>mcgen.CONFIG</i>	25
6.4	DLPOLY FIELD: <i>mcgen.FIELD</i>	25
7	Muudatuste tegemine	25
7.1	Dihedraalfunktsiooni lisamine	25
7.2	Kaugmõju energiafunktsiooni lisamine	25

1 mcgen

mcgen on Metropolis Monte Carlo algoritmil põhinev programm polümeeriahelate juhuslike konfiguratsioonide genereerimiseks. Programmi autoriteks on Heiki Kasemägi, Endel Soolo, Alvo Aabloo ja Josh Thomas. **mcgen** eelkäijateks ja eeskujuks on Mattias Klintenbergi samal algoritmil põhinev C-s kirjutatud programm ja Heiki Kasemägi, Andi Hektori, Mattias Klintenbergi, Alvo Aabloo ja Josh Thomase poolt kirjutatud programm *mcpoly*. **mcgen** on *mcpoly* põhjalik edasiarendus: täielikult on muutunud programmi sisend-väljund; ahelate genereerimise tehnikat on täiustatud ja edasiarendatud nt. kõrvalahelate lisamise näol. Programmi arendatakse Tartu ja Uppsala Ülikoolide koostöös. Programmi arendamisele on finantsiliselt kaasa aidanud Nordic Energy Research (NEFP), Kami Research Foundation (KFS) ja Eesti Teadusfond (ETF).

mcgen ei kuulu hetkel ühegi senise üldiselt teadaoleva litsentsi alla; litsentsina ja programmi kasutustingimustena tuleb vaadelda käesolevas dokumendis toodud tingimusi. Programm on saadaval autorite käest. Teadaolevalt pole programmile hindade kehtestatud. Koodi võib vabalt muudatusi teha, aga need muudatused tuleb saata tagasi autoritele. Programmi edasine levitamine ilma autorite nõusolekuta on keelatud. Programmi, programmi koodi või selle osade kommertskasutamine on ilma autorite kirjaliku nõusolekuta keelatud.

mcgen eesmärk on pakkuda akadeemiliks uurimistööks odavamat alternatiivi analoogsetele kommertsprogrammidele. Kasutajal on otsene juurdepääs koodile ja võimaluse teha muudatusi ning lisada uusi omadusi vastavalt oma vajadustele. Sellised muudatused programmi koodi on igati teretulnud eeldusel, et need lisatud/muudetud kood järgib programmis kasutatavat üldist programmeerimisstiili ja et kõik muudatused oleksid nõuetekohaselt dokumenteeritud.

2 Disclaimer

Ükski **mcgen** autoritest ei garanteeri, et programm ei sisalda vigu. Samuti ei võta autorid endale mingit vastutust programmi kasutamisel tekkida võivate tark- ja riistvarakahjustuste eest. **mcgen** kasutamine ilma tasuta kehtib vaid akadeemilise uurimisöö kohta. Kommertskasutamine on võimalik alles pärast läbirääkimisi autoritega. Kasutajatel pole lubatud levitada programmi kolmandatele osapooltele.

3 Tööpõhimõte

mcgen 's kasutatakse polümeeriahela genereerimiseks Monte Carlo meetodit. Keskne juurdelisatav üksus on *unit*, mis koosneb $1 \dots N$ aatomist. *uniti* iga juurdelisatava aatomi jaoks on vaja teada järgmisi parameetreid:

- kaugus eelmisest aatomist;
- nurk kahe viimase aatomiga;
- dihedraalnurk kolme viimase aatomiga;
- kolme viimase aatomi koordiaadid.

Kaugus, nurk ja dihedraalnurk võivad olla fikseeritud, aga neid saab ka juhuslikult etteantud piirides genereerida. Üldjuhul on soovitatav lasta vähemasti dihedraalnurk programmil juhuarvude põhjal genereerida.

Juurdelisatava *uniti* aatomite genereerimisel kontrollitakse esmalt, et genereeritud aatom ei sattuks mõnele teisele aatomile liiga lähedale või isegi kattuks sellega. Kui midagi sellist juhtub, genereeritakse kohe uued koordiaadid ja teostatakse järjekorrel kontroll kuni uus aatom on piisavalt kaugel teistest aatomitest.

Kui *uniti* kõik aatomid on geomeetriliselt genereeritud, siis rakendub Monte Carlo meetod, mis seisneb lühidalt alljärgnevas:

1. arvutatakse saadud uue konfiguratsiooni koguenergia E ;
2. saadud koguenergia lahutatakse eelmise konfiguratsiooni koguenergiast E_0 : $\Delta E = E - E_0$;
3. kui uue konfiguratsiooni energia on madalam eelmise konfiguratsiooni energiast, s.t. $\Delta E \leq 0$, siis aksepteeritakse saadud konfiguratsioon koheselt ja see saab uue konfiguratsiooni aluseks;

4. kui saadud konfiguratsiooni energia on suurem eelmise konfiguratsiooni energiast, s.t. $\Delta E > 0$, siis võrreldakse Boltzmanni faktorit $\exp(-\Delta E/kT)$ (k - Boltzmanni konstant, T - simulatsiooni temperatuur Kelvinites) juhusliku arvuga vahemikus $0 \dots 1$:

- kui Boltzmanni faktor on suurem sellest juhuarvust, siis aktsepteeritakse uut konfiguratsiooni, aktsepteeritavuse tingimuseks on $\text{rand}(0,1) \leq \exp(-\Delta E/kT)$;
- kui Boltzmanni faktor on väiksem sellest juhuarvust, siis uut konfiguratsiooni ei aktsepteerita ja alustatakse uuesti *uniti* genereerimist eelmise konfiguratsiooni alusel.

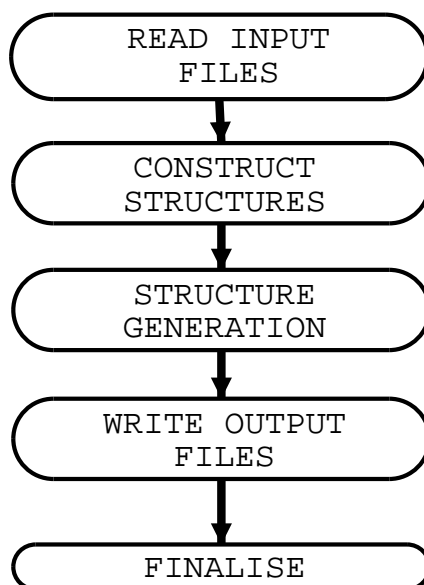
Et ülalkirjedatud meetodil ei leita mitte kõige madalama energiaga konfiguratsioon, vaid esimene lokaalne miinimum, siis teatud tingimustel võib genereerimine joosta ummikusse, s.t. isegi peale suurt proovimiste arvu ei suudeta vajalik hulk *unit*'eid ära genereerida. Sellise olukorra leevendamiseks rakendatakse tagasivõtmise mehhanismi, mille puhul peale fikseeritud proovide arvu hüljatakse etteantud vahemikus juhuslik arv viimaseid *unit*'eid ja alustatakse uuesti genereerimist.

4 Programmi struktuur

4.1 Programmi üldine struktuur

Programmi üldine struktuur on skemaatiliselt joonisel 1.

Figure 1: Programmi struktuur



4.2 Sisendfailide lugemine

Sisendfailide lugemine on skemaatiliselt joonisel 2.

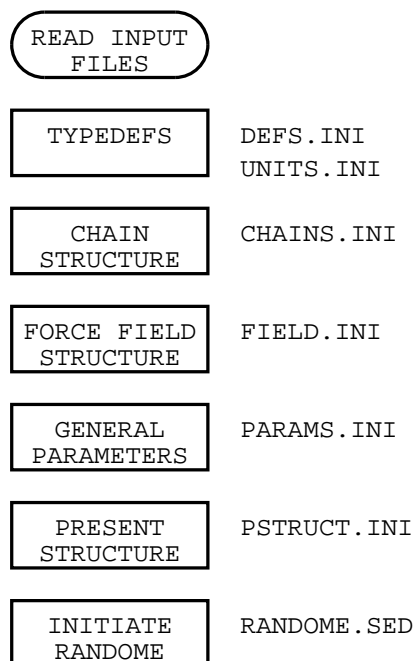
TYPEDEFS-seksioonis loetakse *defs.ini* ja *units.ini* sisendfailidest info aatomite, sidemete, kauguste, nurkade, unitite jne. kohta. Antud info säilitatakse mälus ühekordselt ja on mõeldud kasutamiseks andmebaasina hilisema genereerimise käigus. Iga objekt, mis genereerimise käigus luuakse ja mille struktuur on andmebaasis olemas, saab endale identifikaatori, mis viitab kirjele andmebaasis. Selle abil saab objekti kohta lisainfot, vältides vajadust anda iga loodava objektiga kaasa sedasama korduvat infot ja vähendades kasutatava mälu hulka. *defs.ini* sisendfaili detailne kirjeldus on peatükis 5.2 ja *units.ini* sisendfaili detailne kirjeldus – peatükis 5.3. Antud sisendfailid on üksteise suhtes hierarhilised, s.t. *units.ini* kasutab juba *defs.ini*-s olevat infot.

CHAIN STRUCTURE-seksioonis loetakse *chains.ini* sisendfailist genereeritava süsteemi struktuur: ahelad, nende arv, järjekord, koostis jne. *chains.ini* sisendfail on *TYPEDEF*-seksioonis loetud sisendfailide suhtes hierarhiline,

kasutades neis olevat infot. *chains.ini* sisendfaili detailne kirjeldus on peatükis 5.4.

FORCE FIELD STRUCTURE-seksioonis loetakse kasutatava jõuvälja parameetrid ja struktuur. *field.ini* sisendfail kasutab *defs.ini* isendfailis olevaid definitsioone. *field.ini* sisendfaili detailne kirjeldus on peatükis 5.5.

Figure 2: Sisendfailide lugemine



4.3 Genereeritava struktuuri koostamine sisendinfo põhjal

Sisendfailides *defs.ini*, *units.ini* ja *chains.ini* põhjal koostatakse hierarhiline genereeritava süsteemi struktuur (vt. joonist 3. *defs.ini* põhjal koostatakse aatomitüüpide, sidemepikkuste, sidemete, nurgasidemete, nurkade, dihedraalsidemete ja dihedraalnurkade (ehk kokkuvõtvalt elementide) andmebaasid. Igal elemendil on vastavas andmebaasis oma indeks, mis antakse kaasa genereerimisel loodavale uuele objektile (aatom, unit jne.). Andmebaasid peavad sisaldama kõiki genereerimisel kasutatavate erinevate elementide definitsioone.

Elementide andmebaas sisaldab:

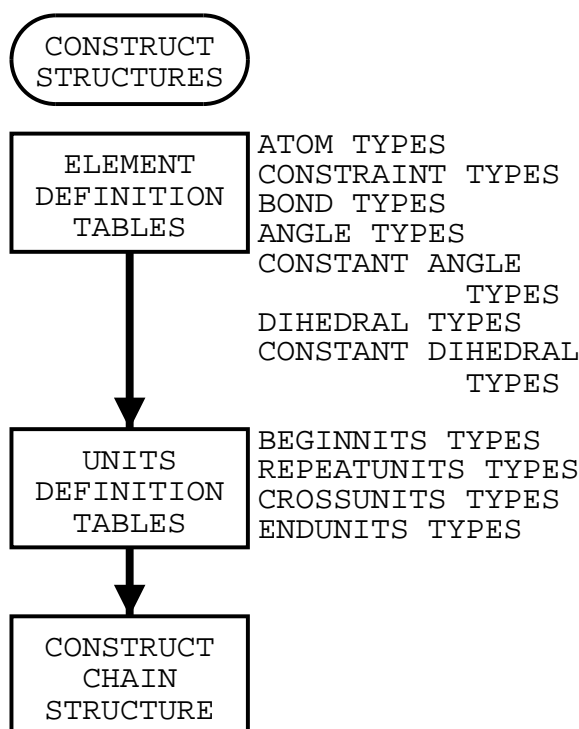
- aatomi tüübikeirjeldust - aatomi nimetus, mass, laeng, raadius;
- sidemepikkuste kirjeldust - kahe aatomi vaheline kaugus ja selle lubatud kõikumine;
- sidemete tüübikeirjeldust - kahe aatomi vahel oleva sideme tüüp ja sidemes kasutatavate parameetrite väärtused;
- nurgasidemete tüübikeirjeldust - kolme aatomi vahel oleva nurgasideme tüüp ja sidemes kasutatavate parameetrite väärtused;
- nurkade tüübikeirjeldust - kolme aatomi vahel olev nurk ja selle lubatud kõikumine;
- dihedraalsidemete tüübikeirjeldust - nelja aatomi vahel oleva dihedraalsideme tüüp ja sidemes kasutatavate parameetrite väärtused;
- dihedraalnurkade tüübikeirjeldust - nelja aatomi vahel oleva dihedraalnurga väärtus.

units.ini põhjal koostatakse kõikvõimalike struktuuris esinevate unitite andmebaas. Siin refereeritakse juba indeksi kaudu elementide andmebaasi. Unit on defineeritud kui omavahel sidemetega seotatud (loogiline) aatomite kogum, mis struktuurile juurde lisatakse, kasutades peatükis 3 kirjeldatud Monte Carlo meetodit. Unitide suurus võib alata ühest aatomist ja olla teoreetiliselt piiramatult suur. Iseärsuste ja asukoha järgi genereeritavas struktuuris on unitid jaotatud nelja kategooriasse:

- BEGINUNITS - üks unit, millega ahel algab; teoreetiliselt võiks neid igas ahelas olla vaid 1; antud unititüüpi kuuluvaid aatomeid saab esitada kahel viisil: kas otse aatomite xyz-koordinaate esitades või lastes aatomid genereerida aatomid juhuslikult juhuslikku kohta simulatsiooniboksis, arvestades aatomite jaoks etteantud vahekaugusi ja nurki;
- REPEATUNITS - unititüüp, mida genereeritakse üksteise järel etteantud arv kordi; selle uniti sees jaotuvad aatomid *backbone*- ja *hydrogen*-aatomteks, millel on tavaliselt tehniliselt erinev aatomi lisamise meetod; *backbone*-aatomid moodustavad antud unititüübi ja kordumisel kogu ahelaosa selgroo, nende aatomitega on seotud ahela võimalikud hargnemised ja *hydrogen*-tüüpi kõrvalaatomid; *hydrogen*-aatom on seotud vaid ühe kindla *backbone*-aatomi ja mitte eühegi teise aatomiga, neid aatomeid saab asendada ahela hargnemisel kõrvalahela *backbone*-aatomiga;
- CROSSUNIT - see on unititüüp, mida kasutatakse hargnemispunktide kirjeldamiseks repeatunititest koosneva ahela sees; teoreetiliselt võik selle struktuur olla sarnane repeatunitiga enne ja pärast hargnemiskohta koos kõrvalahela genereerimiseks vajalike aatomite asendamisega kõrvalahela omadega; analoogselt repeatunititega on aatomid ajagatud *backbone*- ja *hydrogen*-aatomiteks;
- ENDUNIT - üks unit, millega ahel lõpetatakse; tüüpide arv sõltub erinevate ahelatüüpide arvust ja struktuurist; kõik aatomid on unitis ühte tüüpi.

Genereeritav süsteem on hierarhiline: genereeritavad ahelad on üksteise suhtes erineva tähtsusega. Kõige kõrgemal ehk 0-tasemel on peaahelat moodustavad ahelad.

Figure 3: Genereeritava struktuuri koostamine



5 Sisend

5.1 mcgen käsuriida

mcgen käsurea kirjelduse saab lihtsalt programmi nime sisestamisega käsurealt:

```

$ ./mcgen
#EE> ./mcgen: missing command line arguments
#==> try './mcgen --help' for more information
  
```

This indicates that more command line arguments are needed to run the program:

```
./mcgen --help
# Mandatory arguments to long options are mandatory for short options too.
# -d, --defs-file DEFSFILENAME      defs file name
# -p, --params-file PARAMSFILENAME  params file name
# -r, --randome-file RANDEFILNAME   randome file name
# -u, --units-file UNITSFILENAME    units file name
#
# -h, --help                        display this help and exit
# -V, --version                     output version information and exit
```

Käsurea argumendid on põhiliselt mõeldud sisendfailide asukohtade sisestamiseks. Kehtivad nii lühikesed kui ka pikad argumendid. Seejuures on lühikesele või pikale argumendinimele kohustuslik valik kohustuslik ka pikale või lühikesele argumendinimele.

Versiooniinfo annab teada järgmist:

```
$ ./mcgen -V
# Version: mcgen-00.00.04-031127 by cipo@ut.ee
#
# Copyright (C) 2003 Heiki Kasemgi, Endel Soolo
# This is free software; see the source for copying conditions. There is NO
# warranty; not even for MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.
```

5.2 Tüübikirjeldused: *defs.ini*

defs.ini sisaldab kõigi genereerimisel kasutatavate aatomi-, sideme-, kauguse-, nurga- jm. tüüpide kirjeldusi. Faili nimi pole fikseeritud ja see tuleb sisestada **mcgen** käsurealt. Faili formaat on fikseeritud. Kommentaarirea ees peab olema '#'. Muutujatel puudub fikseeritud formaat. Muutujatüübid on string, char[n], integer ja double.

id-väärtuste nummeratsioon algab '1'-st. *atomXid* (X = 1, 2, 3, 4) väärtus '-1' tähendab, et aatomi tüüp pole oluline. Kui *CONSTRAINTS* ja *CANG* sektsioonides väärtus > 0.0 ja *tolerance* > 0.0, siis on väärtuse kõikumine lubatud, kui väärtus > 0.0 ja *tolerance* == 0.0, siis kõikumine pole lubatud, muutuja väärtus peab olema täpne. *CONSTDHEDRALS* sektsioonis tähendab muutuja väärtus '0.0' täielikku juhuväärtuse genereerimist.

Fail on jagatud sektsioonideks. Iga sektsioon algab võtmesõna ja järgnevate kirjade arvuga. Iga kirje algab uuel real võtmesõnaga, millele järgnevad samal real väärtused, nagu kirjeldab järgnev tabel 1.

Table 1: *defs.ini* struktuur

Võtmesõna	Muutuja	Tüüp	Selgitus
DEFINITIONS		string	faili algus
ATOMS		string	aatomite definitsioonide sektsioon
	nr	integer	eri tüüpi aatomite arv
ATOM		string	aatomi kirje
	id	integer	kirje indeks
	label	char[4]	aatomi nimetus
	mass	double	aatomi mass
	charge	double	aatomi laeng
	radius	double	aatomi raadius
CONSTRAINTS		string	kauguste sektsioon
	nr	integer	eri tüüpi kauguste arv
CONST		string	kauguse kirje
	id	integer	kirje indeks
	atom1type	integer	esimese aatomi tüübi indeks ATOMS-sektsioonis
	atom2type	integer	teise aatomi tüübi indeks ATOMS-sektsioonis
	length	double	kaugus ängströmid

Table 1: (...jätkub...)

Võtmesõna	Muutuja	Tüüp	Selgitus
	tolerance	double	lubatud kõikumine ångströmid
BONDS		string	keemiliste sidemete sektsioon
	nr	integer	eri tüüpi sidemete arv
BOND		string	sideme kirje
	id	integer	kirje järjekorranumber
	atom1type	integer	esimese aatomi tüübi indeks ATOMS-sektsioonis
	atom2type	integer	teise aatomi tüübi indeks ATOMS-sektsioonis
	name	char[4]	sideme nimetus
	forcenr	integer	parameetrite arv
	force[1]	double	esimene parameeter
	...		
	force[forcenr]	double	forcenr-is parameeter
ANGLES		string	nurgasidemete sektsioon
	nr	integer	eri tüüpi sidemete arv
ANG		string	nurgasideme kirje
	atom1type	integer	esimese aatomi tüübi indeks ATOMS-sektsioonis
	atom2type	integer	teise aatomi tüübi indeks ATOMS-sektsioonis
	atom3type	integer	kolmanda aatomi tüübi indeks ATOMS-sektsioonis
	name	char[4]	sideme nimetus
	forcenr	integer	parameetrite arv
	force[1]	double	esimene parameeter
	...		
	force[forcenr]	double	forcenr-is parameeter
CONSTANGLES		string	nurkade sektsioon
	nr	integer	eri tüüpi nurkade arv
CANG		string	nurga kirje
	atom1type	integer	esimese aatomi tüübi indeks ATOMS-sektsioonis
	atom2type	integer	teise aatomi tüübi indeks ATOMS-sektsioonis
	atom3type	integer	kolmanda aatomi tüübi indeks ATOMS-sektsioonis
	angle	double	nurga väärtus kraadides
	tolerance	double	lubatud kõikumine kraadides
DIHEDRALS		string	dihedraalsidemete sektsioon
	nr	integer	eri tüüpi sidemete arv
DIH		string	sideme kirje
	id	integer	kirje järjekorranumber
	atom1type	integer	esimese aatomi tüübi indeks ATOMS-sektsioonis
	atom2type	integer	teise aatomi tüübi indeks ATOMS-sektsioonis
	atom3type	integer	kolmanda aatomi tüübi indeks ATOMS-sektsioonis
	atom4type	integer	neljanda aatomi tüübi indeks ATOMS-sektsioonis
	name	char[4]	sideme nimetus
	forcenr	integer	parameetrite arv
	force[1]	double	esimene parameeter
	...		
	force[forcenr]	double	forcenr-is parameeter
CONSTDIHEDRALS		string	dihedraalnurkade sektsioon
	nr	integer	eri tüüpi nurkade arv
CDIH		string	dihedraalnurga kirje
	id	integer	kirje järjekorranumber
	atom1type	integer	esimese aatomi tüübi indeks ATOMS-sektsioonis
	atom2type	integer	teise aatomi tüübi indeks ATOMS-sektsioonis
	atom3type	integer	kolmanda aatomi tüübi indeks ATOMS-sektsioonis
	atom4type	integer	neljanda aatomi tüübi indeks ATOMS-sektsioonis
	angle	double	dihedraalnurk
CLOSE		string	faili lõpp

5.3 uniti kirjeldus: *units.ini*

units.ini sisaldab juurdelisatavate *unit*ite kirjeldusi: aatomid, nurgad, kaugused jne. Faili nimi pole fikseeritud ja see tuleb sisestada *mcgen* käsurealt. Faili formaat on fikseeritud. Kommentaarirea ees peab olema '#'. Muutujatel puudub fikseeritud formaat. Muutujatüübid on string, char[n], integer ja double.

Kõik kirjetes vajalikud tüübiindeksid saadakse *defs.ini* failist. Aatomkirjetes esinev muutuja *method* omab kahte väärtust:

back - aatomi genereerimiseks kasutatakse kolme [järjestikkust] *backbone*-aatomit; aatom seostatakse viimasega neist kolmest;

hydro - aatomi genereerimiseks kasutatakse kolme [järjestikkust] *backbone*-aatomit; aatom seostatakse neist keskmisega.

id-väärtuste nummeratsioon algab '1'-st. *unitXid* ($X = 1, 2, 3$) väärtus '-1' tähendab, et vajalik aatom saadakse unititist, mis vastavalt *chain.ini*-s olevale infole peab olema valmisgenereeritud unit. Väärtus '+1' tähendab, et vajalik aatom saadakse unitist, mis vastavalt *chain.ini*-s olevale infole genereeritakse peale jooksvat unitit. Väärtus '0' tähendab jooksvat unitit.

Fail on jagatud sektsioonideks. Iga sektsioon algab võtmesõna ja sektsioonis olevate kirjeplokkide arvuga. Iga kirjeplokk algab võtmesõna ja plokis olevate kirjete arvuga.

BEGINUNITS-sektsioonis lisandub võtmesõnale ja kirjete arvule aatomite genereerimise meetod, millest sõltub kirje struktuur:

coords - kirjes esitatakse aatomi koordinaadid XYZ-koordinaadistikus;

random - kirjes esitatakse parameetrid aatomite genereerimiseks juhuslikku kohta juhuslikus suunas, kuid kirjetega määratud järjekorras ning üksteisevaheliste kauguste ja nurkadega.

REPEATUNITS- ja *CROSSUNITS*-sektsioonides on *UNIT*-kirjeplokkides aatomite kirjed jaotatud alamplukkideks *BACKBONE* ja *HYDROGENS*, et eraldada loogiliselt ahela selgroogu moodustavaid aatomeid neist aatomeist, mis on sidemetega selgrooaatomitega seotud, kuid millel endal rohkem sidemid pole. Erandiks on siin aatomid, mille külge kinnituvad kõrvalahelad.

units.ini täpsem kirjeldus on toodud tabelis 2.

Table 2: *units.ini* struktuur

Võtmesõna	Muutuja	Tüüp	Selgitus
UNITS		string	faili algus
BEGINUNITS		string	algusunitite sektsioon
UNIT	nr	integer	algusunitite kirjelduste arv
		string	uniti kirjelduse algus
	id	integer	uniti järjekorranumber antud sektsioonis
	atomnr	integer	aatomite arv unitis
	genmeth	string	genereerimise meetod:
genmeth==coords			aatomid esitakse xyz-koordinaatidena
ATOM		string	aatomi kirje
	id	integer	aatomi indeks uniti kirjelduses
	type	integer	aatomi tüüp
	xcoord	double	x-coordinate
	ycoord	double	y-coordinate
	zcoord	double	z-coordinate
genmeth==random			aatomid genereeritakse juhuslikult
ATOM		string	aatomi kirje
	id	integer	aatomi indeks uniti kirjelduses
	type	integer	aatomi tüüp

Table 2: (...jätkub...)

Võtmesõna	Muutuja	Tüüp	Selgitus
	constid	integer	kauguse tüüp
	cangid	integer	nurga tüüp
	cdihid	integer	dihedraali tüüp
	method	string	genereerimise meetod:
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
REPEATUITS		string	korduvunitite sektsioon
UNIT	nr	integer	korduvunitite kirjelduste arv
		string	uniti kirjeldus
BACKBONE	id	integer	uniti index antud sektsioonis
	atomnr	integer	aatomite arv unitis: <i>backbone+hydrogens</i>
		string	<i>backbone</i> -aatomite alamsektsioon
ATOM	atomnr	integer	<i>backbone</i> -aatomite arv
		string	aatomi kirje
	id	integer	aatomi index
	type	integer	aatomi tüüp
	constid	integer	kauguse tüüp
	cangid	integer	nurga tüüp
	cdihid	integer	dihedraalnurga tüüp
	method	string	genereerimise meetod
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
HYDROGENS	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
		string	<i>hydrogen</i> -aatomite alamsektsioon
	atomnr	integer	<i>hydrogen</i> -aatomite arv
ATOM		string	aatomi kirje
	id	integer	aatomi index
	type	integer	aatomi tüüp
	constid	integer	kauguse tüüp
	cangid	integer	nurga tüüp
	cdihid	integer	dihedraalnurga tüüp
	method	string	genereerimise meetod
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
CROSSUNITS	atom2id	integer	teise aatomi indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
		string	ristunitite sektsioon
	nr	integer	ristunitite kirjelduste arv
UNIT		string	uniti kirjeldus
	id	integer	uniti indeks antud sektsioonis
BACKBONE	atomnr	inteer	aatomite arv unitis
		string	<i>backbone</i> -aatomite alamsektsioon
	atomnr	integer	<i>backbone</i> -aatomite arv
ATOM		string	aatomi kirje
	id	integer	aatomi index
	type	integer	aatomi tüüp
	constid	integer	kauguse tüüp

Table 2: (...jätkub...)

Võtmesõna	Muutuja	Tüüp	Selgitus
HYDROGENS	cangid	integer	nurga tüüp
	cdihid	integer	dihedraalnurga tüüp
	method	string	genereerimise meetod
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
		string	hydrogen-aatomite alamseksioon
	atomnr	integer	hydrogen-aatomite arv
ATOM		string	aatomi kirje
	id	integer	aatomi index
	type	integer	aatomi tüüp
	constid	integer	kauguse tüüp
	cangid	integer	nurga tüüp
	cdihid	integer	dihedraalnurga tüüp
	method	string	genereerimise meetod
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
ENDUNITS		string	lõpu-unitite sektsioon
	nr	integer	lõpu-unitite kirjelduste arv
UNIT		string	uniti kirjeldus
	id	integer	uniti indeks antud sektsioonis
ATOM	atomnr	integer	aatomite arv unitis
		string	aatomi kirje
	id	integer	aatomi index
	type	integer	aatomi tüüp
	constid	integer	kauguse tüüp
	cangid	integer	nurga tüüp
	cdihid	integer	dihedraalnurga tüüp
	method	string	genereerimise meetod
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
CLOSE		string	faili lõpp

5.4 Ahelate struktuur: *chains.ini*

chains.ini sisaldab genereeritava struktuuri kirjeldust: unitite ja neist moodustatavate ahelate järgnevus, tihedus, determineeritus jne. Faili nimi pole fikseeritud ja see tuleb sisestada **mcgen** käsurealt. Faili formaat on fikseeritud. Kommentaarirea ees peab olema '#'. Muutujatel puudub fikseeritud formaat. Muutujatüübid on string, char[n], integer ja double.

Kõik kirjetes vajalikud tüübiindeksid saadakse *defs.ini* failist. *CHAIN*-kirjes algavad *chainid*-, *levelid*-, *unitid*-muutjate väärtused '1'-st.

Fail on jagatud sektsioonideks. Neist esimesed kaks, *CHAINS* ja *CHAIN unittype==begin* korral, esinevad vaid

ühe korra, sest *CHAINS* on faili alguskirje ja igal genereeritaval struktuuril saab olla vaid üks algusunit.

Tavaline ahelakirje (*unitttype==repeat* korral) koosneb esmalt alguskirjest *CHAIN*, millele järgneb *CONNECTFIRST*-kirje. See kirje sisaldab infot järgneva ahelaosa ühendamiseks eelnenud ahelaosaga. Sellele kirjele järgneb *CROSSES*-kirje, mis sisaldab infot võimalike kõrvalahelate kohta lisatavas ahelaosas: võimalikud ristunitite tüübid, mille seast võib valida, kõrvalahelate lisamise algoritmi ja arvu jne. Kui *crossnr > 0*, siis peab järgnevaid kirjepaare olema *crossnr* tükki. *CONNECTCROSS* ja *CONNECT2CROSS* käivad paaris sama tüüpi hargnemise kohta. *CONNECTCROSS* kirjeldab, kuidas tuleb ristunit ühendada eelneva ahelaga, *CONNECT2CROSS* aga kirjeldab, kuidas tuleb pärast katkestust ahelat jätkates see ristunitiga ühendada.

NB: *CONNECT2CROSS* ei kirjelda kõrvalahela ühendamist ristunitiga. Seda funktsiooni täidab uues, kõrvalahela jaoks mõeldud *CHAIN*-seksioonis *CONNECTFIRST*-kirje, mida peab olema 1 iga võimaliku ristuniti tüübi jaoks. *crossstype*-muutuja väärtus peab olema sama, mis võimaliku ristuniti tüüp, mille külge ahel tuleks ühendada.

Kui *unitttype==end*, siis puuduvad ahela seksioonist *CONNECTCROSS*- ja *CONNECT2CROSS*-kirjed.

chains.ini kirjeldus on tabelis 3.

Table 3: *chains.ini* struktuur

Võtmesõna	Muutuja	Tüüp	Selgitus
<i>CHAINS</i>		string	faili algus
	levelnr	integer	ahelatasemete arv
	chainnr	integer	ahelakirjete arv
<i>unitttype==begin</i> <i>CHAIN</i>			ahela antud osa unitid on <i>BEGINUNITS</i> -tüüpi
		string	ahelakirje
	chainid	integer	ahela indeks
	levelid	integer	tasemeindeks
	unitnr	integer	genereeritavate unitite arv
	unitttype	string	unitite tüüp
	unitid	integer	uniti indeks antud tüübiglassis
<i>unitttype==repeat</i> <i>CHAIN</i>			ahela antud osa unitid on <i>REPEATUNITS</i> -tüüpi
		string	ahelakirje
	chainid	integer	ahela indeks
	levelid	integer	tesemeindeks
	unitnr	integer	genereeritavate unitite arv
	unitttype	string	unitite tüüp
	unitid	integer	uniti indeks antud tüübiglassis
<i>CONNECTFIRST</i>	cfnr	integer	järgnevate <i>CONNECTFIRST</i> -kirjete arv
		string	esimese uniti genereerimiseks vajalikud parameetrid
	id	integer	kirje index
	typeid	integer	selle uniti tüübiindeks, mille külge ahel pannakse
	constid	integer	kauguse tüüp
	cangid	integer	nurga tüüp
	cdihid	integer	dihedraali tüüp
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
		string	<i>CROSSUNITITS</i> -ite valimine
	crossnr	integer	valitavate <i>CROSSUNITITS</i> -ite tüüpide arv
	crossstype[1]	integer	esimene <i>CROSSUNITITS</i> -i tüüp
	...		
<i>CROSSES</i>			

Table 3: (...jätkub...)

Võtmesõna	Muutuja	Tüüp	Selgitus
CONNECTCROSS	crosstype[<i>crossnr</i>]	integer	<i>crosstype</i> -s <i>CROSSUNITS</i> -ite tüüp
	repeat	integer	lisatavate kõrvalahelate arv
	spacing	integer	lisatavate kõrvalahelate kaugus üksteisest (ühik: unit)
	tolerance	integer	lisatava kõrvalahela lubatud nihe (ühik: unit)
		string	ristuniti ühendamise; jäetakse ära, kui <i>crossnr</i> == 0
	id	integer	ristühenduse indeks
	crosstypeid	integer	ühendatava <i>CROSSUNITS</i> tüüp
	constid	integer	kauguse tüüp
	cangid	integer	nurga tüüp
	cdihid	integer	dihedraali tüüp
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhtline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhtline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhtline indeks
	atom3id	integer	kolmanda aatomi indeks
CONNECT2CROSS		string	ahela jätku ühendamise ristunitiga
	id	integer	ühenduse index
	crosstypeid	integer	<i>CROSSUNITS</i> -i tüüp, mille külge ühendatakse
	constid	integer	kauguse tüüp
	cangid	integer	nurga tüüp
	cdihid	integer	dihedraali tüüp
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhtline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhtline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhtline indeks
	atom3id	integer	kolmanda aatomi indeks
CHAIN	unitttype==end		ahela antud osa unitid on <i>ENDUNITS</i> -tüüpi
		string	ahelakirje
	chainid	integer	ahela indeks
	levelid	integer	tasemeindeks
	unitnr	integer	genereeritavate unitite arv
	unitttype	string	unitite tüüp
	unitid	integer	uniti indeks antud tüübiklassis
	cfnr	integer	järgnevate <i>CONNECTFIRST</i> -kirjete arv
		string	enduniti ühendamiseks vajalikud parameetrid
	id	integer	kirje index
	typeid	integer	selle uniti tüübiindeks, mille külge ahel pannakse
	constid	integer	kauguse tüüp
	cangid	integer	nurga tüüp
	cdihid	integer	dihedraali tüüp
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhtline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhtline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks

Table 3: (...jätkub...)

Võtmesõna	Muutuja	Tüüp	Selgitus
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
CLOSE		string	faili lõpp

5.5 Jõuväli: *field.ini*

field.ini fail sisaldab genereeritava struktuuri jõuvälja kirjeldust, et genereerimise käigus struktuuri energiat arvutada. Faili nimi pole fikseeritud ja see tuleb sisestada **mcgen** käsurealt. Faili formaat on fikseeritud. Kommentaarirea ees peab olema '#'. Muutujatel puudub fikseeritud formaat. Muutujatüübid on string, char[n], integer ja double.

Kõik kirjetes vajalikud tüübiindeksid saadakse *defs.ini*, *chains.ini*- ja *units.ini*-failidest. *id*-väärtused algavad '1'-st.

chainXid ($X = 1, 2, 3, 4$) väärtus on ahela indeks vastavalt *chains.ini*-failile. Kui *chainXid* == 0, siis on tegemist jooksva ahelaga. Selle muutuja väärtus ei saa olla negatiivne.

unitXid väärtus on suhteline, märkides jooksvat unitit '0' korral, eelmist või eelmisi uniteid '-1, -2,...' korral ning järgmist või järgmisi uniteid nullist suuremate positiivsete väärtuste korral.

atomXid väärtus on aatomi indeks vastava uniti kirjelduses *units.ini* failis.

Fail on jagatud sektsioonideks ja järgib oma struktuurilt *chains.ini*-faili. Tegelikult ei saa faili mõistlikul viisil enne *chains.ini*-faili loomist koostada.

Fail on jagatud sektsioonideks lisatavate unitite tüübi järgi. Kui uniti tüübiks on *BEGINUNITS*, siis koosneb *FIELD*-kirje plokk vaid sidemetest, mis kirjeldavad aatomitevahelisi energetilisi lähimõjuseid vaid ühe uniti piires. Selliste sidemeplokkide tähisteks siin ja edaspidi on *INxxxx*, kus *xxxx* = *BONDS*, *ANGLES* või *DIHEDRALS*.

Kui uniti tüübiks on *REPEATUNITS*, siis lisanduvad *FIELD*-kirjesse muutujad *cfnr* ja *cunr*. Esimene neist väljendab *CFxxxx* kirjeplokkide arvu vastavalt *chains.ini* kirjeldatud võimalikele ühendusviisidele. Teine muutuja väljendab võimalike ristunitite arvu vastavalt *chains.ini*-failile. *INxxxx*-kirjete plokkide võib olla rohkem kui üks vastavalt *chains.ini* failile. Lisaks seostele unitite sees ja samatüübilistele unititele sisaldavad *INxxxx*-plokkid kirjeid ahela ühenduskohtade jaoks.

INxxxx-plokkidele järgnevad *CFxxxx*-plokkid, mis kirjeldavad seoseid vastavalt *CONNECTFIRST*-kirjetele *chains.ini*-failis. *CFxxxx*-plokkidele järgnevad *CCxxxx*-plokkid, mis kirjeldavad seoseid vastavalt *CONNECTCROSS*-kirjetele *chains.ini*-failis. *CCxxxx*-kirjetele järgnevad *C2Cxxxx*-kirjed, mis kirjeldavad seoseid vastavalt *CONNECT2CROSS*-kirjetele *chains.ini*-failis.

Kui *unittyp*==*end*, siis sisaldab *FIELD*-kirjete plokk *INxxxx*-kirje(te) ja *CFxxxx*-kirje(te) plokk(e)i ja *FIELD*-kirje ei sisalda *cunr*-muutujat.

field.ini kirjeldus on tabelis 4.

Table 4: *field.ini* struktuur

Võtmesõna	Muutuja	Tüüp	Selgitus
FIELDS		string	faili algus
	fieldnr	integer	<i>FIELD</i> -kirjete arv
	levelnr	integer	tasemete arv
unittyp==begin			
FIELD		string	jõuväljakirje

Table 4: (...jät kub...)

Võttesõna	Muutuja	Tüüp	Selgitus	
	fieldid	integer	kirjeindeks	
	levelid	integer	tasmeindeks	
	unittype	unitite tüüp		
	INBONDS		string	unitit sees ja samatüübiliste unitite vahel olevad sidemed
BOND	nr	integer	järgnevate <i>BOND</i> -kirjete arv	
	id	integer	sideme indeks	
	typeid	integer	sideme tüübiindeks	
	chain1id	integer	esimese aatomi ahela indeks	
	unit1id	integer	esimese aatomi uniti suhteline indeks	
	atom1id	integer	esimese aatomi indeks	
	chain2id	integer	teise aatomi ahela indeks	
	unit2id	integer	teise aatomi uniti suhteline indeks	
INANGLES	atom2id	integer	teise aatomi indeks	
		string	unitit sees ja samatüübiliste unitite vahel olevad nurgasidemed	
	nr	integer	järgnevate <i>ANG</i> -kirjete arv	
	id	integer	sideme indeks	
ANG	typeid	integer	sideme tüübiindeks	
	chain1id	integer	esimese aatomi ahela indeks	
	unit1id	integer	esimese aatomi uniti suhteline indeks	
	atom1id	integer	esimese aatomi indeks	
	chain2id	integer	teise aatomi ahela indeks	
	unit2id	integer	teise aatomi uniti suhteline indeks	
	atom2id	integer	teise aatomi indeks	
	chain3id	integer	kolmanda aatomi ahela indeks	
	unit3id	integer	kolmanda aatomi uniti suhteline indeks	
	atom3id	integer	kolmanda aatomi indeks	
	INDIHEDRALS		string	unitit sees ja samatüübiliste unitite vahel olevad dihedraalsidemed
	DIH	nr	integer	järgnevate <i>DIH</i> -kirjete arv
id		integer	sideme indeks	
typeid		integer	sideme tüübiindeks	
chain1id		integer	esimese aatomi ahela indeks	
unit1id		integer	esimese aatomi uniti suhteline indeks	
atom1id		integer	esimese aatomi indeks	
chain2id		integer	teise aatomi ahela indeks	
unit2id		integer	teise aatomi uniti suhteline indeks	
atom2id		integer	teise aatomi indeks	
chain3id		integer	kolmanda aatomi ahela indeks	
unit3id		integer	kolmanda aatomi uniti suhteline indeks	
atom3id		integer	kolmanda aatomi indeks	
chain4id		integer	neljanda aatomi ahela indeks	
unit4id		integer	neljanda aatomi uniti suhteline indeks	
atom4id		integer	neljanda aatomi indeks	
unittype==repeat				
FIELD		string	jõuväljakirje	
	fieldid	integer	kirjeindeks	
	levelid	integer	tasmeindeks	
	unittype	unitite tüüp		
	cfnr	integer	järgnevate <i>INxxxx</i> -kirjeplokkide arv	
	cunr	integer	<i>INxxxx</i> -kirjeplokkidele järgnevate <i>Cxxxx</i> -kirjeplokkide arv	
INBONDS		string	unitit sees ja ahela alguse ühendamiseks vajalikud sidemed	
BOND	nr	integer	järgnevate <i>BOND</i> -kirjete arv	
	id	integer	sideme indeks	
	typeid	integer	sideme tüübiindeks	

Table 4: (...jätkub...)

Võtmesõna	Muutuja	Tüüp	Selgitus
INANGLES	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
		string	unitit sees ja ahela alguse ühendamiseks vajalikud nurgasidemed
ANG	nr	integer	järgnevate <i>ANG</i> -kirjete arv
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
		string	unitit sees ja ahela alguse ühendamiseks vajalikud dihedraalsidemed
DIH	nr	integer	järgnevate <i>DIH</i> -kirjete arv
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
	chain4id	integer	neljanda aatomi ahela indeks
	unit4id	integer	neljanda aatomi uniti suhteline indeks
	atom4id	integer	neljanda aatomi indeks
CFBONDS		string	sidemed eelmis(t)e ahela(te)ga
	nr	integer	järgnevate <i>BOND</i> -kirjete arv
BOND	connectid	integer	<i>CONNECT</i> -kirje indeks
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
		string	nurgasidemed eelmis(t)e ahela(te)ga
CFANGLES	nr	integer	järgnevate <i>ANG</i> -kirjete arv
	connectid	integer	<i>CONNECT</i> -kirje indeks
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks

Table 4: (...jätkub...)

Võtmesõna	Muutuja	Tüüp	Selgitus
CFDIHEDRALS	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
DIH		string	dihedraalsidemed eelmis(t)e ahela(te)ga
	nr	integer	järgnevate <i>DIH</i> -kirjete arv
	connectid	integer	<i>CONNECT</i> -kirje indeks
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
	chain4id	integer	neljanda aatomi ahela indeks
	unit4id	integer	neljanda aatomi uniti suhteline indeks
	atom4id	integer	neljanda aatomi indeks
CCBONDS		string	sidemed ristuniti sees ning unitiga, mille külge ristunit ühendatakse
BOND	nr	integer	järgnevate <i>BOND</i> -kirjete arv
	crosstypeid	integer	ühendatava ristuniti tüübiindeks
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
CCANGLES		string	nurgasidemed ristunitit sees ning unitiga, mille külge ristunit ühendatakse
ANG	nr	integer	järgnevate <i>ANG</i> -kirjete arv
	crosstypeid	integer	ühendatava ristuniti tüübiindeks
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
CCDIHEDRALS	atom3id	integer	kolmanda aatomi indeks
		string	dihedraalsidemed ristunitit sees ing unitiga, mille külge ristunit ühendatakse
DIH	nr	integer	järgnevate <i>DIH</i> -kirjete arv
	crosstypeid	integer	ühendatava ristuniti tüübiindeks
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks

Table 4: (...jätkub...)

Võttesõna	Muutuja	Tüüp	Selgitus
C2CBONDS	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
	chain4id	integer	neljanda aatomi ahela indeks
	unit4id	integer	neljanda aatomi uniti suhteline indeks
	atom4id	integer	neljanda aatomi indeks
BOND	nr	string	sidemed ristunitiga ahela jätkamisel
	crosstypeid	integer	järgnevate <i>BOND</i> -kirjete arv
	id	integer	ühendatava ristuniti tüübiindeks
C2CANGLES	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	nr	string	nurgasidemed ristunitiga ahela jätkamisel
	crosstypeid	integer	järgnevate <i>ANG</i> -kirjete arv
	id	integer	ühendatava ristuniti tüübiindeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
C2CDIHEDRALS	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
	nr	string	dihedraalsidemed ristunitiga ahela jätkamisel
	crosstypeid	integer	järgnevate <i>DIH</i> -kirjete arv
	id	integer	ühendatava ristuniti tüübiindeks
	typeid	integer	sideme tüübiindeks
DIH	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
	chain4id	integer	neljanda aatomi ahela indeks
	unit4id	integer	neljanda aatomi uniti suhteline indeks
	atom4id	integer	neljanda aatomi indeks
unittype==end			
FIELD		string	jõuväljakirje

Table 4: (...jätkub...)

Võttesõna	Muutuja	Tüüp	Selgitus
INBONDS	fieldid	integer	kirjeindeks
	levelid	integer	tasmeindeks
	unitttype	unitite tüüp	
	cfnr	integer	järgnevate <i>INxxxx</i> -kirjeplokkide arv
BOND	nr	string	unitit sees ja ahela alguse ühendamiseks vajalikud sidemed
	nr	integer	järgnevate <i>BOND</i> -kirjete arv
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
		string	unitit sees ja ahela alguse ühendamiseks vajalikud nurgasidemed
ANG	nr	integer	järgnevate <i>ANG</i> -kirjete arv
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
INDIHEDRALS	atom3id	integer	kolmanda aatomi indeks
		string	unitit sees ja ahela alguse ühendamiseks vajalikud dihedraalsidemed
	nr	integer	järgnevate <i>DIH</i> -kirjete arv
DIH	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
	chain4id	integer	neljanda aatomi ahela indeks
	unit4id	integer	neljanda aatomi uniti suhteline indeks
	atom4id	integer	neljanda aatomi indeks
CFBONDS		string	sidemed eelmis(t)e ahela(te)ga
	nr	integer	järgnevate <i>BOND</i> -kirjete arv
BOND	connectid	integer	<i>CONNECT</i> -kirje indeks
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks

Table 4: (...jätkub...)

Võtmesõna	Muutuja	Tüüp	Selgitus
CFANGLES	nr	string	nurgasidemed eelmis(t)e ahela(te)ga
	connectid	integer	järgnevate <i>ANG</i> -kirjete arv
ANG	id	integer	<i>CONNECT</i> -kirje indeks
	typeid	integer	sideme indeks
	chain1id	integer	sideme tüübiindeks
	unit1id	integer	esimese aatomi ahela indeks
	atom1id	integer	esimese aatomi uniti suhteline indeks
	chain2id	integer	esimese aatomi indeks
	unit2id	integer	teise aatomi ahela indeks
	atom2id	integer	teise aatomi uniti suhteline indeks
	chain3id	integer	teise aatomi indeks
	unit3id	integer	kolmanda aatomi ahela indeks
	atom3id	integer	kolmanda aatomi uniti suhteline indeks
		integer	kolmanda aatomi indeks
		string	dihedraalsidemed eelmis(t)e ahela(te)ga
	nr	integer	järgnevate <i>DIH</i> -kirjete arv
DIH	connectid	integer	<i>CONNECT</i> -kirje indeks
	id	integer	sideme indeks
	typeid	integer	sideme tüübiindeks
	chain1id	integer	esimese aatomi ahela indeks
	unit1id	integer	esimese aatomi uniti suhteline indeks
	atom1id	integer	esimese aatomi indeks
	chain2id	integer	teise aatomi ahela indeks
	unit2id	integer	teise aatomi uniti suhteline indeks
	atom2id	integer	teise aatomi indeks
	chain3id	integer	kolmanda aatomi ahela indeks
	unit3id	integer	kolmanda aatomi uniti suhteline indeks
	atom3id	integer	kolmanda aatomi indeks
	chain4id	integer	neljanda aatomi ahela indeks
	unit4id	integer	neljanda aatomi uniti suhteline indeks
	atom4id	integer	neljanda aatomi indeks
CLOSE		string	faili lõpp

so far so good...do not care about the rest at the moment

5.6 Üldised sisendparameetrid: *params.ini*

params.ini sisaldab ahela genereerimiseks vajalikke abiparameetreid. Faili nimi pole fikseeritud ja see tuleb sisetada **mcgen** käsurealt. *params.ini* faili loeb klass PARAMS. Üks *params.ini* näide on toodud allpool:

```
Amorphous poly(ethylene oxide)
CELL1 20.0 0.0 0.0
CELL2 0.0 20.0 0.0
CELL3 0.0 0.0 20.0
IMCON 1
B_LNG 1.3
MAXEN 1000000
MAXTRY 10000000000
MAXTRYUNIT 500
GEOMCONSTRNR 0
TEMPERATURE 300
CUTOFF 10.0 0.0001
RANDOM 2 10000
CLOSE
```

Faili formaat on fikseeritud. Võtmesõnad peavad olema suurtähtedega. Tühje ridasid ei tohi võtmesõnade vahel esineda. Muutujatel puudub fikseeritud formaat. Muutujatüübid on char[8], integer, long ja double.

header - ahela nimetus

char[80]

CELL1 CELL[0] CELL[1] CELL[2] - ühikraku maatriksi esimene rida

fikseeritud võtmesõna : double : double : double

CELL[0] - ühikraku maatriksi esimese rea esimene element

CELL[1] - ühikraku maatriksi esimese rea teine element

CELL[2] - ühikraku maatriksi esimese rea kolmas element

CELL2 CELL[3] CELL[4] CELL[5] - ühikraku maatriksi teine rida

fikseeritud võtmesõna : double : double : double

CELL[3] - ühikraku maatriksi teise rea esimene element

CELL[4] - ühikraku maatriksi teise rea teine element

CELL[5] - ühikraku maatriksi teise rea kolmas element

CELL3 CELL[6] CELL[7] CELL[8] - ühikraku maatriksi kolmas rida

fikseeritud võtmesõna : double : double : double

CELL[6] - ühikraku maatriksi kolmanda rea esimene element

CELL[7] - ühikraku maatriksi kolmanda rea teine element

CELL[8] - ühikraku maatriksi kolmanda rea kolmas element

IMCON IMCON - kujutise tüüp

fikseeritud võtmesõna : integer

IMCON=0 - ääritingimused puuduvad

IMCON=1 - standardsed kuubilised ääritingimused

IMCON=2 - ortorombilised ääritingimused

IMCON=3 - 'parallelepiped' ääritingimused

IMCON=4 - lõigatud oktaeedrilised ääritingimused

IMCON=5 - rombilised dodekaheedrilised ääritingimused

IMCON=6 - x-y parallelogramm-ääritingimused; z-suunas perioodilisus puudub

B_LNG B_LNG - keemilise sideme baaspikkus

fikseeritud võtmesõna : double

MAXEN MAXEN - ahela maksimaalne energia

fikseeritud võtmesõna : long

MAXTRY MAXTRY - proovide arv ahela kohta

fikseeritud võtmesõna : long

MAXTRYUNIT MAXTRYUNIT - proovide arv ühe *uniti* kohta

fikseeritud võtmesõna : long

GEOMCONSTRNR GEOMCONSTRNR - erinevate geomeetriliste piirangute arv

fikseeritud võtmesõna : integer

Nüüd järgneb 'GEOMCONSTRNR' arv piirangute kirjeldusi. Igat piirangut kirjeldab alljärgnev kirje:

GEOM GMTYPE GMPAR1 GMPAR2 GMPAR3 GMPAR4 GMPAR5 GMPAR6 GMPAR7

fikseeritud võtmesõna : integer : double : double : double : double : double : double : double

GMTYPE - geomeetrilise piirangu tüüp; olemasolevad tüübid:

0 - piirangud puuduvad

1 - sfääriline piirang: määratud sfääri sisse ahelat ei genereerita

2 - kiht: määratud kihist väljapoole ahelat ei genereerita

9 - kasutaja defineeritud piirang

GMPAR1 - esimene piiranguparameeter

GMPAR2 - teine piiranguparameeter

GMPAR3 - kolmas piiranguparameeter

GMPAR4 - neljas piiranguparameeter

GMPAR5 - viies piiranguparameeter

GMPAR6 - kuues piiranguparameeter

GMPAR7 - seitsmes piiranguparameeter

Kui geomeetrilisi piiranguid pole, siis peab 'GEOMCONSTR' väärtus olema '0' ja peale seda võtmesõna ei tohi olla ühtegi 'GEOM' kirjet.

GMTYPE=1 on sfääriline piirang. GMPAR1, GMPAR2 ja GMPAR3 on selle sfääri keskpunkti x-, y- ja z-koordinaadid. Ahelat selle sfääri sisse ei genereerita. GMPAR4 on selle sfääri raadius. Teised parameetrid GMPAR* (*=5,6,7) peavad olema '0.0'.

GMTYPE=2 on teatud paksusega kiht xy-tasandil, mille sisse ahel genereeritakse. GMPAR1 on kihi alumise välis-tasandi z-koordinaat ja GMPAR2 on ülemise välis-tasandi z-koordinaat. Teised parameetrid GMPAR* (*=3,4,5,6,7) peavad olema '0.0'.

GMTYPE=9 on kasutaja defineeritud piirang. Piirangu kood peab olema funktsioonis

```
user_geom_constr(int imcon, double *unitcell, double x, double y, double z, double par1, double par2, double par3, double par4, double par5, double par6, double par7)
```

failis *userfunc.cpp*.

Sisend:

int imcon - kujutistüüp 'IMCON'

double *unitcell - ühikrakk 'CELL'

double x - jooksva aatomi x-koordinaat

double y - jooksva aatomi y-koordinaat

double z - jooksva aatomi z-koordinaat

double par1 - piiranguparameeter 'GMPAR1'

double par2 - piiranguparameeter 'GMPAR2'

double par3 - piiranguparameeter 'GMPAR3'

double par4 - piiranguparameeter 'GMPAR4'

double par5 - piiranguparameeter 'GMPAR5'

double par6 - piiranguparameeter 'GMPAR6'

double par7 - piiranguparameeter 'GMPAR7'

TEMPERATURE TEMPERATURE - simulatsiooni temperatuur Kelvini skaalas

fikseeritud võtmesõna : double

CUTOFF CUTTOFFDIST CUTOFFSTEP - kaugmõjuinteraktsioonide lõikekaugus

fikseeritud võtmesõna : double : double

CUTTOFFDIST - kaugmõjuinteraktsioonide lõikekaugus

CUTOFFSTEP - kaugmõjuinteraktsioonide lõikesamm

RANDOME RANDVAL RANDSTART - tagasivõtu parameetrid

fikseeritud võtmesõna : long : long

RANDVAL - tagasivõetavate *unit*ite maksimaalne arv

RANDSTART - proovide arv *unit*i kohta, millest alates rakendatakse tagasivõtmist

CLOSE - faili lõpp

fikseeritud võtmesõna

5.7 Kasutaja defineeritavad geomeetrilised piirangud: *userfunc.cpp*

userfunc.cpp fail sisaldab funktsiooni 'user_geom_constr' kasutaja poolt defineeritavate geomeetriliste piirangute jaoks. Faili nimi on fikseeritud. Täienduste kasutamiseks tuleb **mcgen** uuesti kompileerida. Seda funktsiooni kutsutakse välja klassis COORDS. Esialgne kasutaja defineeritud geomeetriliste piirangute funktsioon on failis *userfunc.cpp* alljärgnev:

```
#include <math.h>
#include <stdio.h>
#include <stdlib.h>
#include "image.h"

int user_geom_constr(int imcon, double *unitcell, double x,\\
double y, double z, double par1, double par2, double par3,\\
double par4, double par5, double par6, double par7)
{
int tmp;
tmp=0;

return tmp;
};
```

userfunc.h sisaldab funktsiooni päist.

Funktsioon on kasutatav, kui 'GEOMCONSTR' muutuja failis *params.ini* on suurem kui '0' ja vähemalt ühe 'GEOM' kirje muutuja 'GMTYPE' on '9'.

Sisend:

int imcon - kujutistüüp 'IMCON'
double *unitcell - ühikrakk 'CELL'
double x - jooksva aatomi x-koordinaat
double y - jooksva aatomi y-koordinaat
double z - jooksva aatomi z-koordinaat
double par1 - piiranguparameeter 'GMPAR1'
double par2 - piiranguparameeter 'GMPAR2'
double par3 - piiranguparameeter 'GMPAR3'
double par4 - piiranguparameeter 'GMPAR4'
double par5 - piiranguparameeter 'GMPAR5'
double par6 - piiranguparameeter 'GMPAR6'
double par7 - piiranguparameeter 'GMPAR7'

Väljund:

int user_geom_constr - piirangu staatus
0 - piiranguid pole

1 - olemasolev piirang jooksva aatomi jaoks

Kui kasutaja tahab lisada rohkem kui ühe sisseehitatutest erineva geomeetrilise piirangu, siis peavad need kõik olema selles samas 'user_geom_constr' funktsioonis. Kasutaja defineeritud piirangutele saab *params.ini* faili kaudu edastada vaid 7 parameetrit. Ülejäänud parameetrid tuleb otse funktsiooni sisse kirjutada. Nende igakordne muutmine nõuab *mcgen* ümberkompileerimist.

5.8 Juhuarvude generaator: *randome.sed*

randome.sed fail sisaldab kahte iduarvu juhuarvude generaatori jaoks. Faili nimi pole fikseeritud ja tuleb sisestada *mcgen* käsurealt. Seda faili loeb klass RNDM. Üks *randome.sed* faili näide on alljärgnev"

```
732858166 521611455
```

Faili formaat on fikseeritud. Muutujatel pole fikseeritud formaati. Muutujate tüübiks on long.

```
seed1 seed2
```

```
long : long
```

```
seed1 - esimene iduarv juhuarvude generaatorile
```

```
seed2 - teine iduarv juhuarvude generaatorile
```

mcgen avab selle faili ja loeb juhuarvude generaatori iduarve ühe korra simulatsiooni jooksul. Simulatsiooni lõpul kirjutatakse iduarvude hetkeväärtused *randome.sed* faili.

6 Väljund

6.1 Logifail: *mcgen.log*

mcgen.log sisaldab infot genereerimise kulgemise kohta. Genereerimise jooksul on faili võimalik online jälgimine 'tail -f mcgen.log' käsuga (*nix süsteemides).

mcgen.log sisaldab alguses alljärgneva üldise info parameetrite kohta:

- programmi versioon
- programmi autorid ja kontaktaadressid
- genereeritavate *unit*ite arv
- aatomite arv *unit*is
- ühikraku maatriks
- kujutistüüp
- geomeetriliste piirangute arv ja nende tüübid, kui piiranguid esineb
- simulatsiooni temperatuur
- kaugmõjuinteraktsioonide löikekaugus ja –samm
- tagasivõtu parameetrid
- esialgsed iduarvud juhuarvude generaatorile
- väljundfailide nimed

Seejärel järgneb N (N on määramatu arv ja sõltb konkreetsest situatsioonist) arv kirjeid *unit*ite staatuse kohta. Iga kirje sisaldab jooksva *unit*i numbrit (nummeratsioon algab '0'-st); genereerimise staatuse ja järgmise genereeritava *unit*i numbrit, s.t. kas asutakse genereerima järgmist *unit*it või minnakse mingil põhjusel 1. . . n sammu tagasi; ka tuuakse ära tehtud proovide arv jooksva *unit*i jaoks ja proovide koguarv. Eduka genereerimise korral sisaldab kirje veel keemiliste, nurga- ja dihedraalsidemete ning kaugmõju interaktsioonide energia nii jooksva *unit*i jaoks kui ka summaarselt; uue konfiguratsiooni energia erinevuse eelnevaga ja koguenergia.

Generatsiooni lõppedes korratakse sama infot parameetrite kohta, mis faili alguseski, vaid iduarvude jaoks esitatakse hetkeväärtused, mis kirjutatakse ka *randome.sed* faili. Kõige lõpus on energiatega lõppväärtused ja väljundfailide nimed.

6.2 XYZ: *mcgen.xyz*

mcgen.xyz sisaldav genereeritud aatomite koordinaate xyz-formaadis:

- aatomite arv
- päis
- aatomi_nimetus x y z
- ...
- aatomi_nimetus x y z

Fail on loetav programmidega 'rasmol' ja 'molmol'.

6.3 DLPOLY CONFIG: *mcgen.CONFIG*

mcgen.CONFIG sisaldab genereeritud aatomite koordinaate DLPOLY CONFIG-formaadis.

6.4 DLPOLY FIELD: *mcgen.FIELD*

mcgen.FIELD sisaldab genereeritud ahelale vastava jõuvälja parameetreid DLPOLY FIELD-formaadis. Fail pole koheselt DLPOLY sisendina kasutatav, sest vajab muudatusi vastavalt juhtnööridele, muudele tingimustele või vähemasti nende juhtnööride kustutamist.

7 Muudatuste tegemine

7.1 Dihedraalfunktsiooni lisamine

Dihedraalnurga energiat arvutab klassi CHAIN meetod `torsion_energy`:

```
double CHAIN::torsion\_energy(int \_potkey, double tau, double a0,  
double a1, double a2, double a3, double a4, double a5, double a6)
```

int _potkey on energiafunktsiooni indeks; *double tau* on dihedraalnurk. *double a0 ... double a6* on sisendid dihedraalfunktsiooni parameetrite jaoks. Meetodi sees toimub õige funktsiooni valik if-lausega *_potkey* järgi. Seejärel tuleb klassi FIELD konstruktoris suurendada konstanti *NDIHKEYS* ühe võrra ja kopeerida dihedraalfunktsiooni lühinimi (kuni 4 sümbolit) massiivi *dihkeysname*, nt.:

```
strcpy(dihkeysname[1], "sylv\0");
```

Lõpetusks tuleb klassi COORDS meetodi `retr_dihfieldstr` lisada FIELD-faili jaoks uue potentsiaali vastava formaadi ja parameetrite arvuga string.

7.2 Kaugmõju energiafunktsiooni lisamine

Kaugmõju energiat arvutab klassi CHAIN meetod `vdw_energy`:

```
double CHAIN::vdw_energy(double ln1, int indx, double b1, double b2,  
double b3, double b4, double b5)
```

double ln1 on aatomite vahekaugus; *int indx* on energiafunktsiooni indeks. *double b1 ... double b5* on sisendid energiafunktsiooni parameetritele. Meetodi sees toimub õige funktsiooni valik if-lausega *indx* järgi. Seejärel tulene suurendada konstanti *POTNR* klassi VDW header-faili *vdw.h* alguses ning lisada potentsiaali nimetus (kuni 4 sümbolit) klassi VDW konstruktoris massiivi *potname*, nt.:

```
strcpy(potname[1], "dibu\0");
```