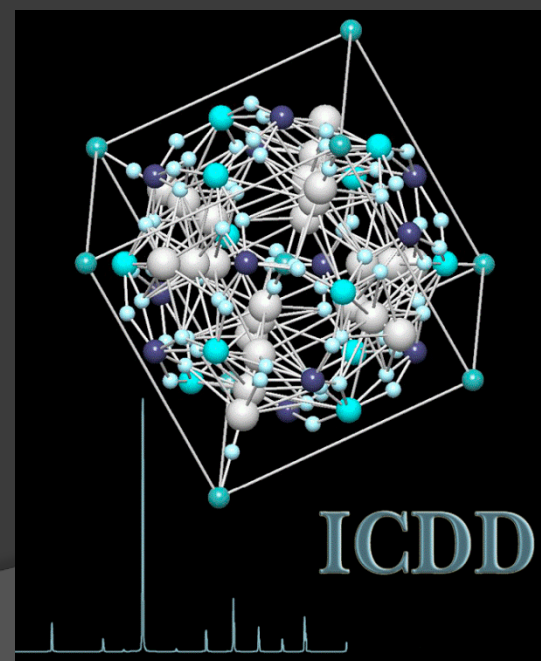


**MATERIAL IDENTIFICATION:
THE DESCRIPTIVE STATISTICS OF PHASE IDENTIFICATION**

T.G. Fawcett, V. Bosnic, S. N. Kabbekodu, F. Needham, J. R. Blanton, D. M. Crane,
C.E. Crowder

International Centre for Diffraction Data



For Pharmaceutical Analyses - Using Descriptive Statistics

Why does Phase Identification work ?

What is the confidence in a correct identification ?

What is the confidence in a trace or minor phase analysis ?

What is the confidence in setting a specification ?

Tools Developed for use with Powder Diffraction File

Relational Database
JAVA Interfaces
48 Searches
72 Display Fields



Data Mining

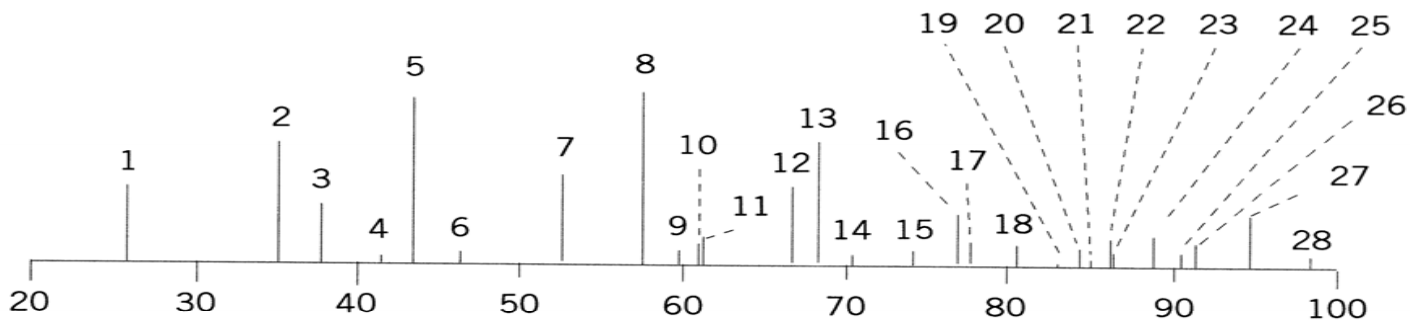
Statistics and Graphic Display Programs

Phase Identification

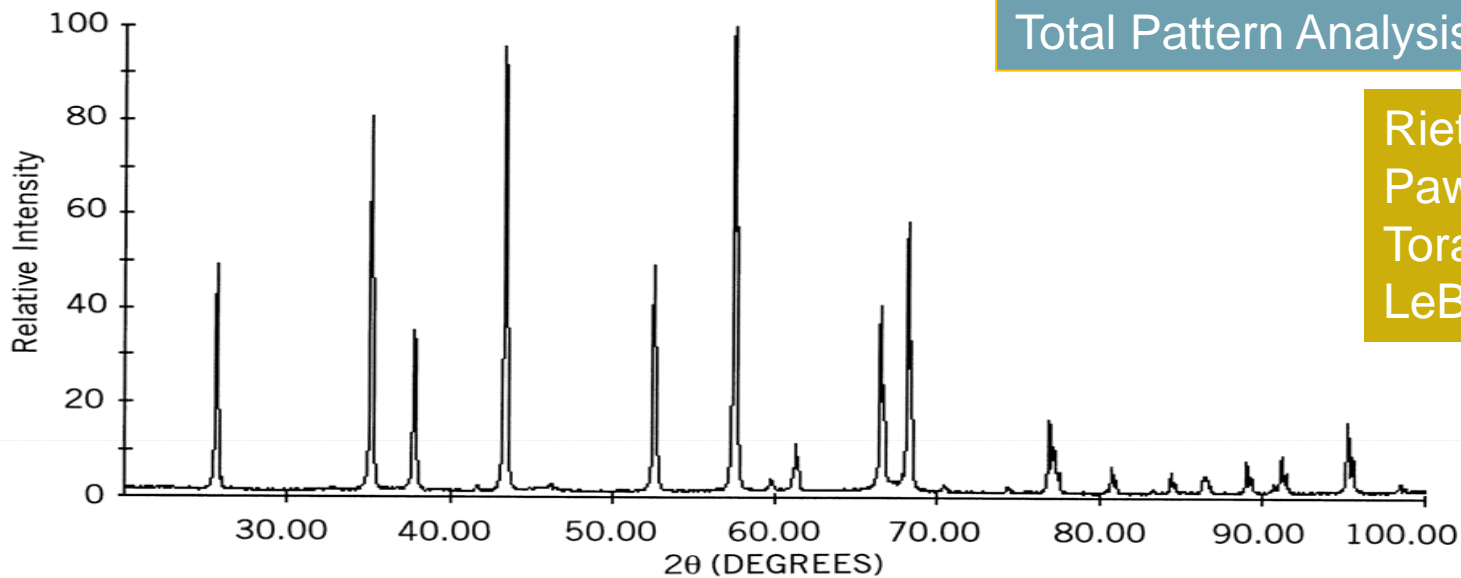
Experimental pattern Positions and Intensities

Hanawalt & Rinn
1936

Reduced data d, I pairs

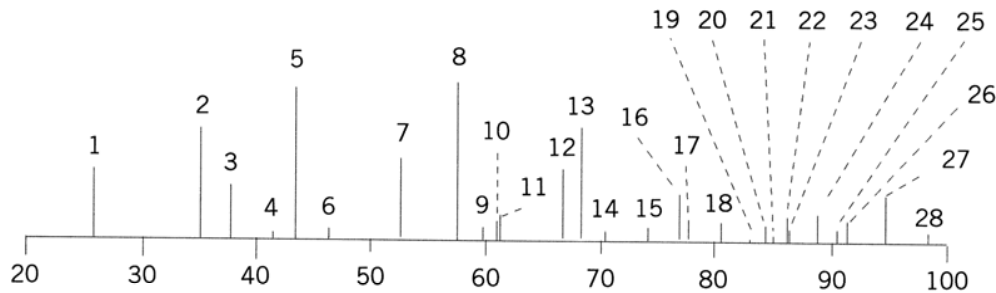


Total Pattern Analysis



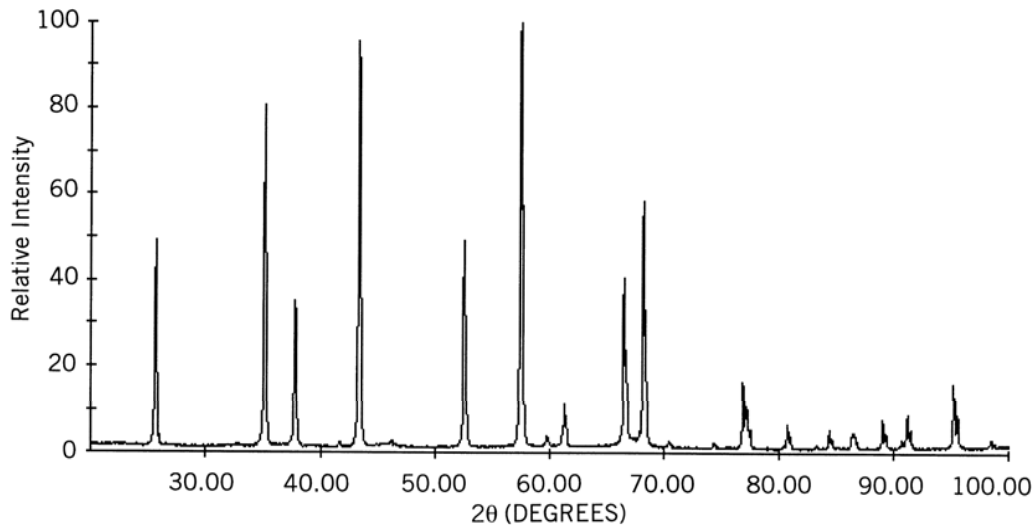
Rietveld 1967
Pawley 1981
Toraya 1986
LeBail 1988

Phase Identification



Use of d,l pairs

Where are the peaks ?



Use of full patterns

**Where are the peaks ?
Where aren't the peaks ?**

Lattice Parameters Determination from Powder Diffraction Data: Results from a Round Robin Project

Norberto Masciocchi

Dipartimento di Chimica Strutturale e Stereochimica Inorganica, Università di Milano, via Venezian 21, 20133 Milano, Italy
Gilberto Artioli

Dipartimento di Scienze della Terra, Università di Milano, via Botticelli 23, 20133 Milano, Italy

Reprinted from *Powder Diffraction*, **11**(3), 253–258 (September 1996).

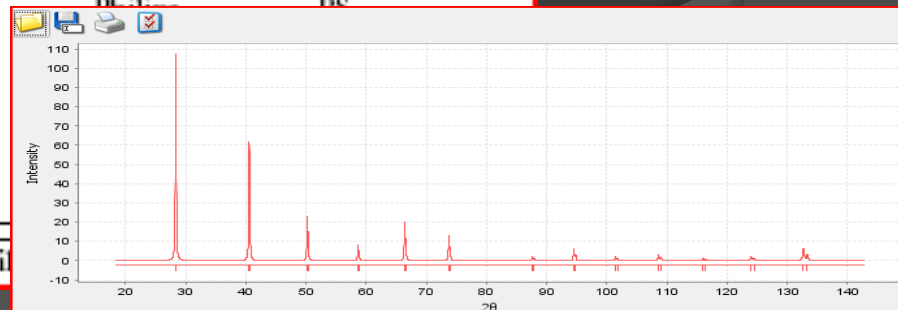
Table 2. Synoptic collection for the 15 independent runs on KCl.

Run No.	N_{peaks}	$2\theta_0, ^\circ$	$\langle \Delta 2\theta \rangle, ^\circ$	$a_0, \text{Å}$	Instrument	Pk^a
1	12	0.0195	0.0055	6.29291(8)	Philips	PS
2	12	-0.0121	0.0028	6.29210(8)	Siemens	PF
3	8	0.1727	0.0091	6.29895(29)	Philips	PM
4	11	0.0604	0.0083	6.29368(13)	Siemens	PS
5	12	0.0580	0.0029	6.29370(4)	Siemens	PS
6	12	0.0090	0.0065	6.29296(11)	Siemens	PS
7	11	-0.0022	0.0101	6.29302(16)	Ital Structures	PF
8	12	0.0344	0.0034	6.29201(7)	Rigaku	PS
9	12	0.0394	0.0020	6.29181(3)	Rigaku	PS
10	11	0.0400	0.0023	6.29198(7)	Rigaku	PS
11	12	-0.0017	0.0040	6.29371(8)	Philips	PS
12	12	-0.0032	0.0071	6.29357(8)	Philips	PS
13	12	0.0050	0.0071	6.29232(15)	Philips	PS
14	12	0.0001	0.0068	6.29217(12)	Philips	PS
15	10	0.1999	0.0038	6.29582(11)	Philips	PS

Simple Pattern

No Standards but
Zero pt shift

Eliminated obvious
Poor data



^aPeak location method: PM = top count; PS = peak search; PF = profile

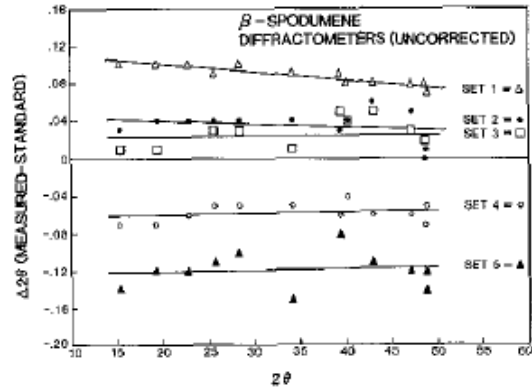


Fig. 4

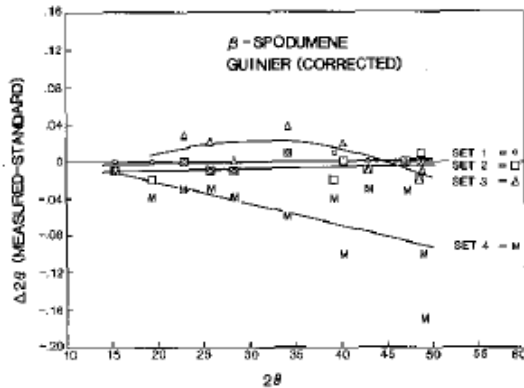


Fig. 5

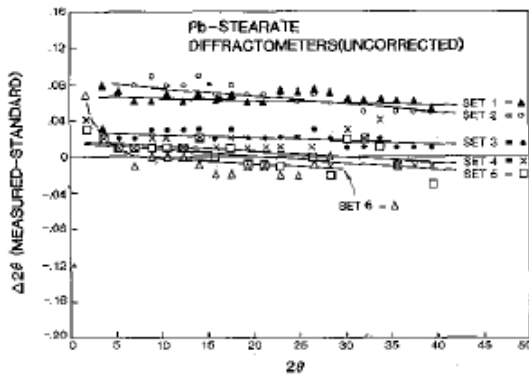


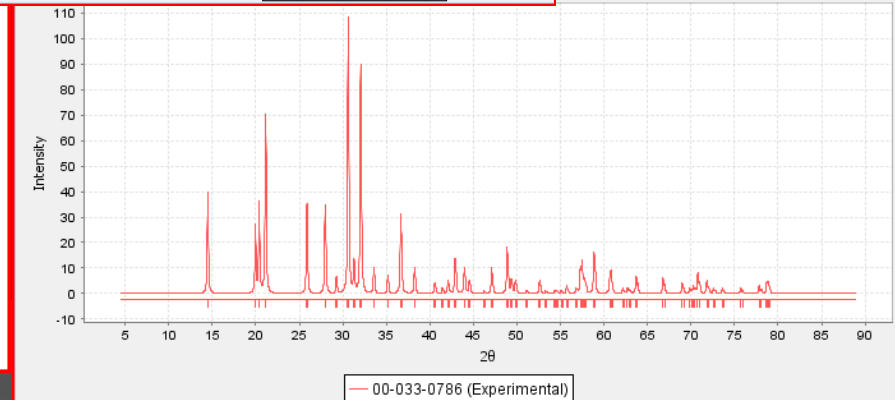
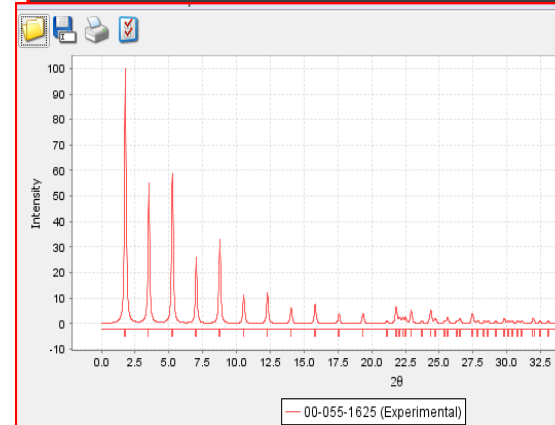
Fig. 6

Results of a Round Robin Study of Systematic Errors Found in Routine X-ray Diffraction Raw Data

Walter N. Schreiner and Tim Fawcett, Vol. 28, pp. 309-314 (1984).

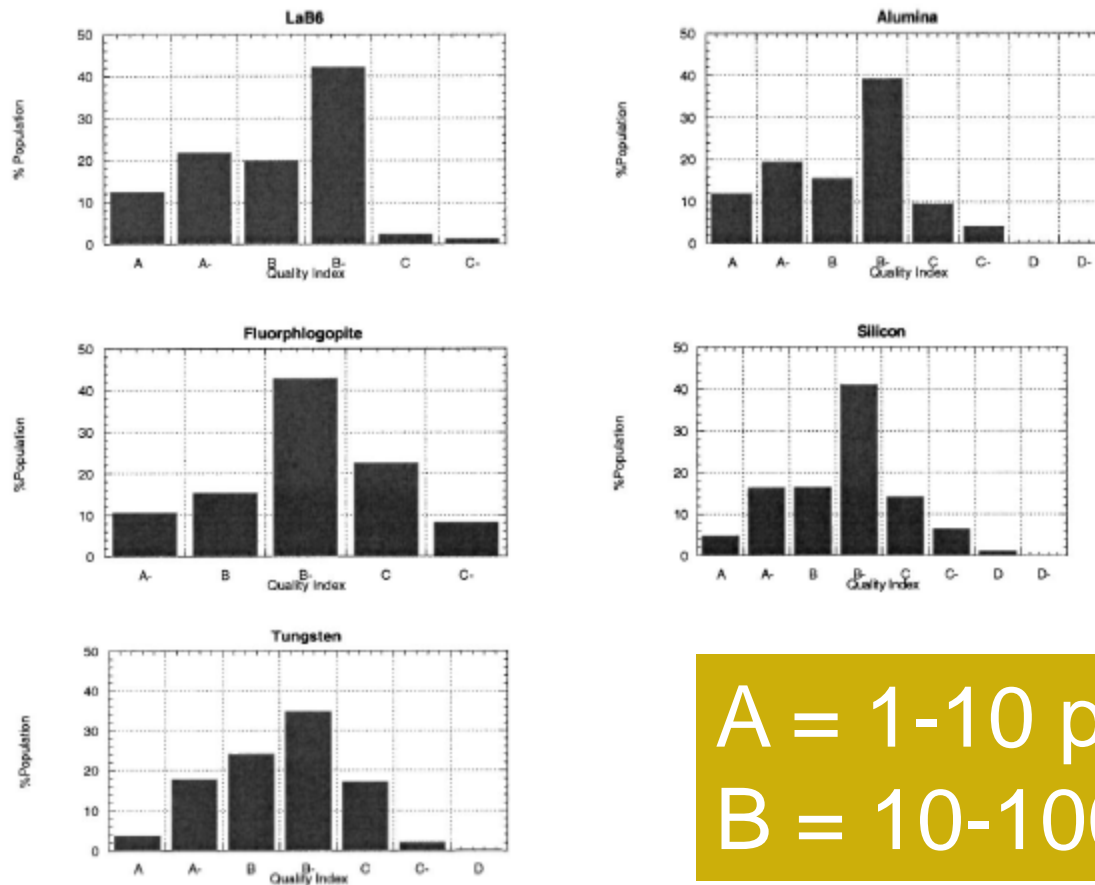
Errors in delta two theta up to 0.08 are common

Standards help significantly



Use of NIST SRM's in Powder Diffraction

- Q (cell esd/cell) evaluation in the PDF



A = 1-10 ppm
B = 10-100 ppm

Figure 6. Quality Index population (%) for data sets using various external and internal standards: (a) lanthanum hexaboride, 170 data sets, (b) alumina, 1048 data sets, (c) fluorophlogopite, 240 data sets, (d) silicon, 9655 data sets, and (e) tungsten, 540 data sets.

Intensity

JCPDS — International Centre for Diffraction Data
Instrument Data Collection Task Group

Intensity Round Robin Report

By Ron Jenkins

JCPDS-ICDD, Swarthmore, Pennsylvania 19081, U.S.A.

Walter N. Schreiner

Philips Laboratories, Briarcliff Manor, New York 10510, U.S.A.

Table 8.

ZnO/C_aCO₃ Intensity and Precision Multi User, Multi Lab

LINE	FDS DATA SETS			VDS DATA SETS		
	AVG	SIGMA	SIG (%)	AVG	SIGMA	SIG (%)
Z/100	71.0	0.0	0	71.0	0.0	0
Z/002	51.3	2.9	6	53.7	7.1	13
Z/110	42.7	5.1	12	65.8	9.0	14
Z/103	34.0	6.4	19	58.1	4.3	7
Z/112	29.0	4.0	14	51.6	4.0	8
Z/213	10.0	1.4	14	25.3	5.5	22
C/012	6.0	3.7	61	6.9	3.3	48
C/104	77.9	35.4	45	103.8	42.1	41
C/113	13.9	7.9	57	20.7	9.1	44
C/202	11.6	6.7	57	18.6	8.1	44
C/116	14.2	6.6	46	26.5	11.1	42
C/314	2.5	1.9	76	7.0	3.1	44

22 Participants

Multiple Tables
indicating poor
precision in intensity
measurements

Conclusion

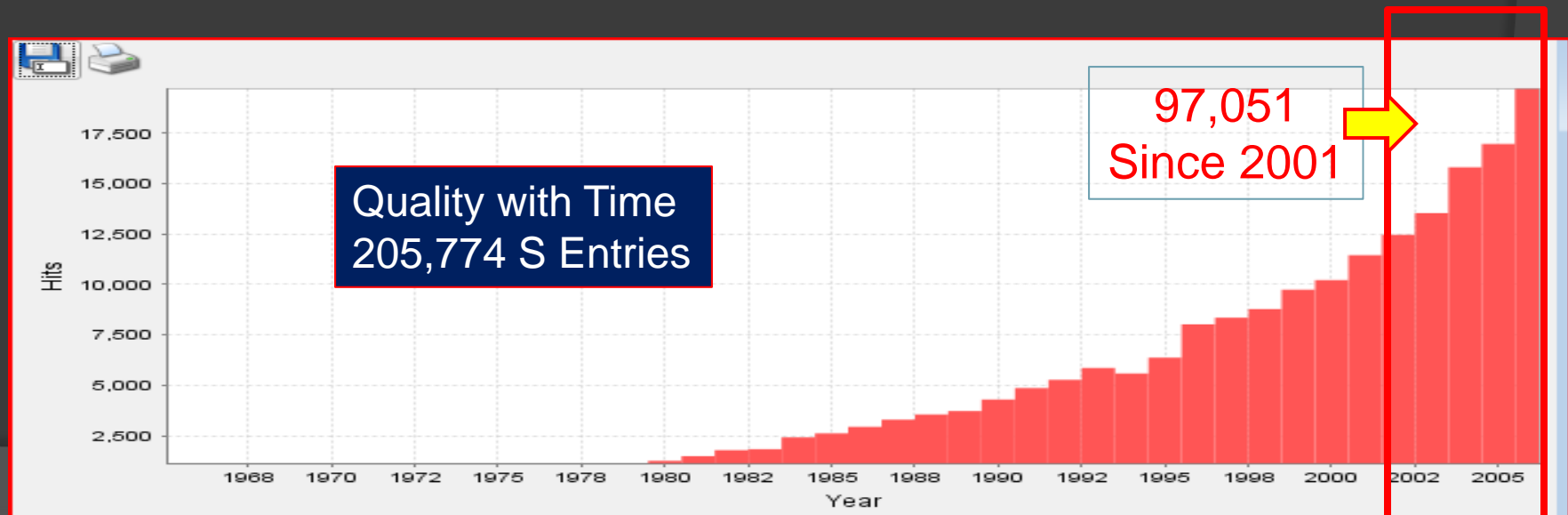
Phase Identification works because of the precision and accuracy in measuring d-spacings

No calibration

1-10 parts per thousand

With standard calibration

1-100 parts per million



Focus on Precision and Accuracy in d-spacings

Careful Specimen Preparation
Careful Instrument Calibration
Use of certified standards

Phase Identification – Uses a team approach

Required of the user

Good specimen preparation (particle statistics, preferred orientation)

Appropriate particle size for good counting statistics

Standardize and calibrate the instrument (Certified Standards)

Internal standard for the most important work

Use instrument settings appropriate for what your trying to measure

Done by the ICDD – Effective database

Data are standardized

Reference quality is measured and evaluated

A range of tools are available for supplemental analytical observations
(searches, indexes, physical properties, structural classifications)

Done by the vendor

Rigorous validated methods of Search/Match

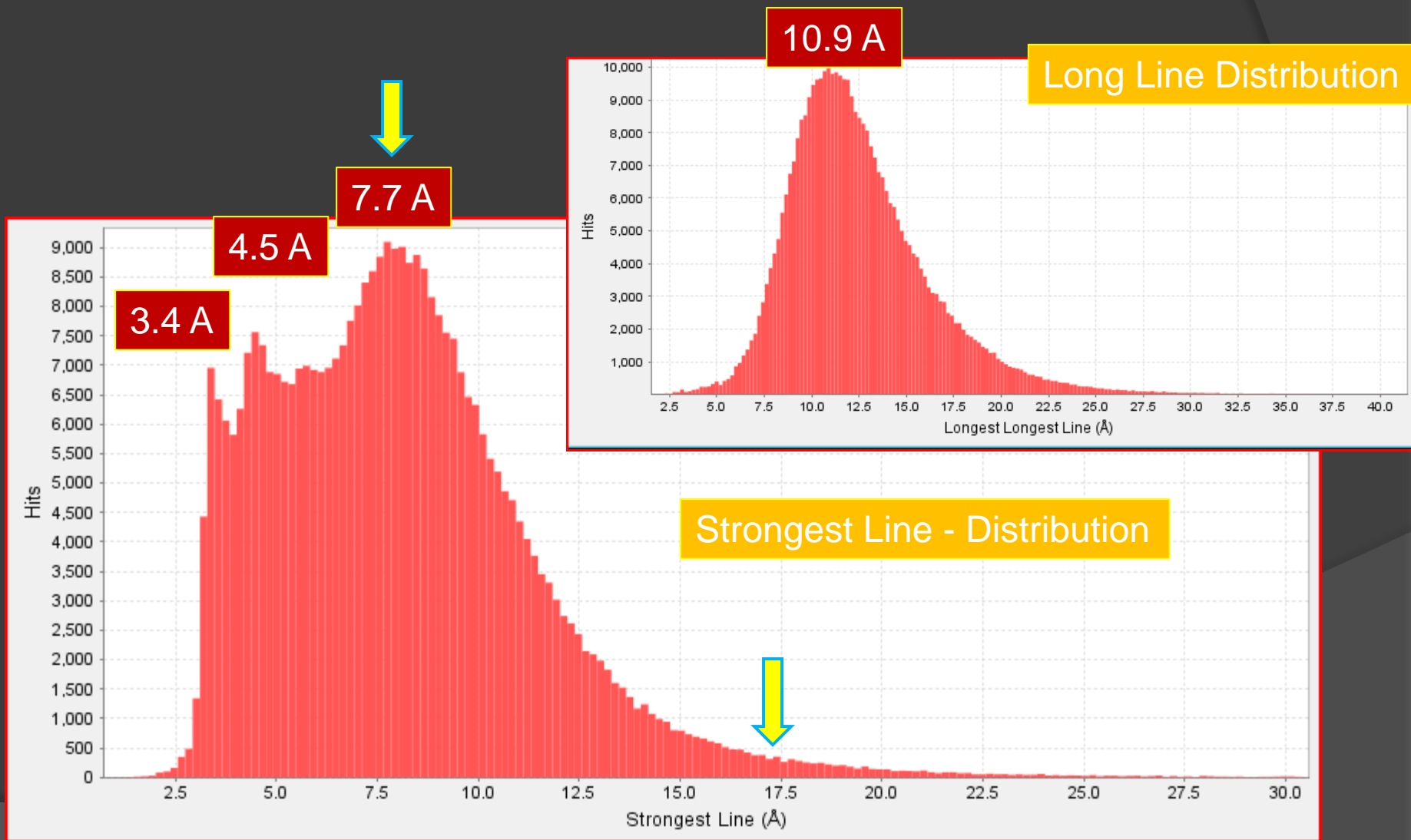
Batch processing for high speed, cluster tools for looking at large data sets

Availability of high resolution optics and highly efficient detectors

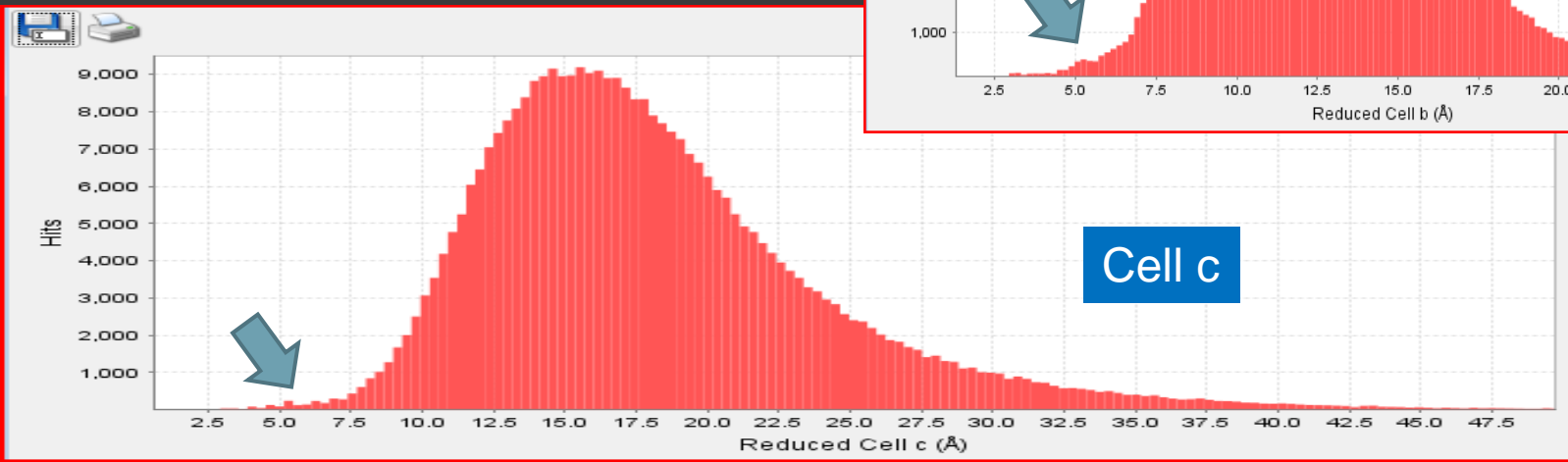
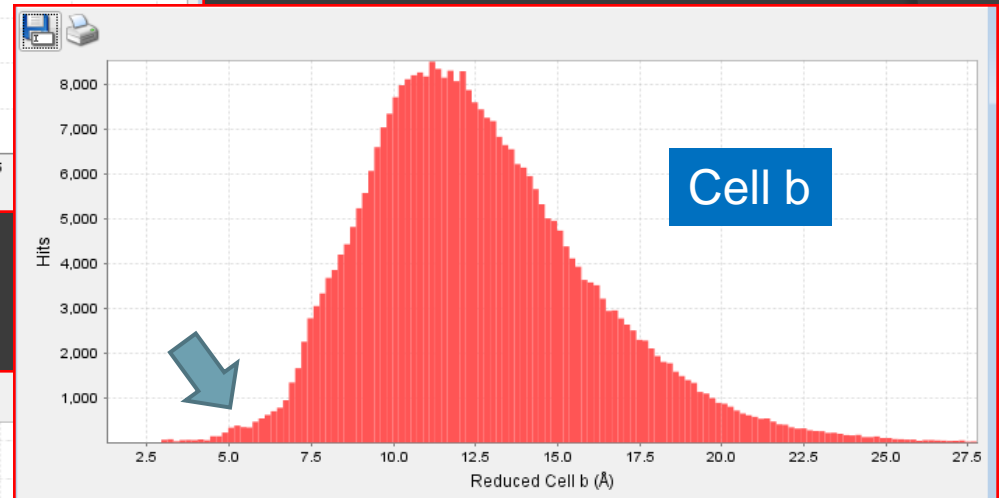
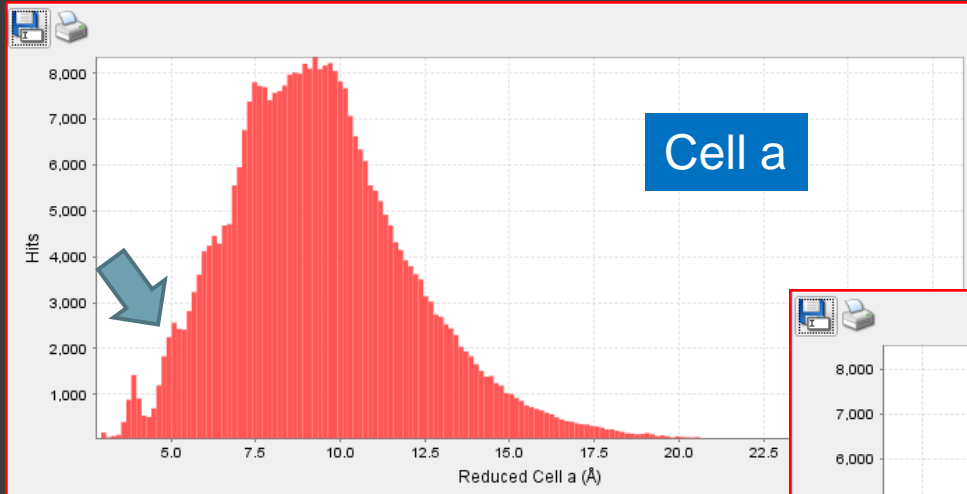
Tools for assisting in calibration and instrument standardization

PDF-4/Organics Release 2009

370,844 Data Sets



Reduced cell edges



Accuracy and Precision

Strong Line Search(d1,d2,d3) –
PDF-4/Organics

7. (1) 198,402 out of 370,844
(1.11 Million d-spacings)

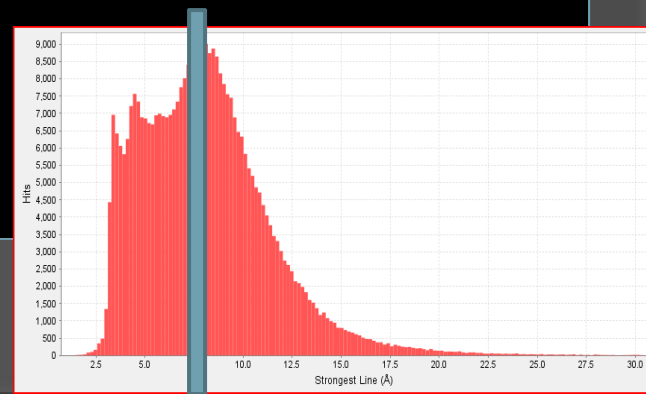
7.2 (0.1) A 27,559

7.22 (0.01) A 2,839

7.222 (0.001) A 291

7.2222 (0.0001) A 33

7.22222 (0.00001) A 1



Olanzapine Form X

Title: **Olanzapine crystal modification**

Document Type and Number: United States Patent 6740753

Abstract: A novel crystal form the pharmaceutical compound olanzapine, processes for its preparation and its pharmaceutical use are disclosed.

11.05 (1) 88,231

11.05 (0.1) 11,186

11.05 (0.05) 5,267

11.05 (0.01) 1,055

Based on d1-d3
search of 1.1 Million
d-spacings

Entire database has
> 66 Million d-spacings

Claim 3 - What does about 11.05
mean ? (1% or 0.1 d-spacing)

Using multiple observations

A look at 7.22 (0.01)

All Data	2,839 Entries
And “blue”	76 Entries
A pharmaceutical	6 Entries
Contains Cl	678 Entries
Contains N	1,775 Entries
Contains C between 7 and 13 %	52 Entries

PDF-4/Organics – 370,884 Entries

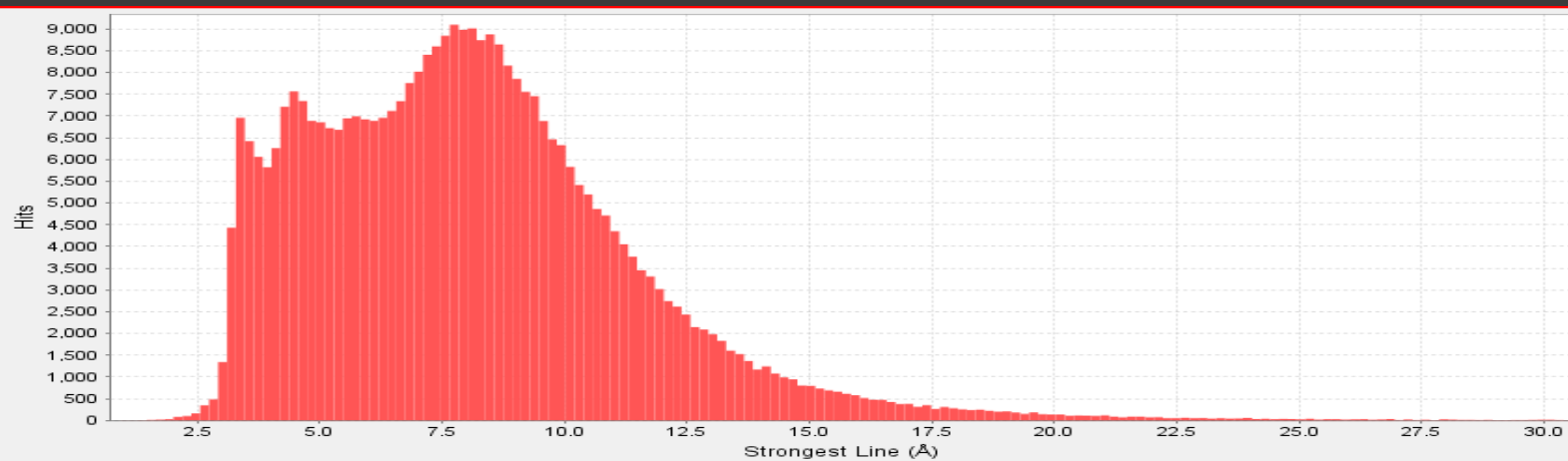
3.22 (0.01)

2,080

7.22 (0.01)

2,839

8 Entries with both



Combined Results

3.22(1), 7.22(1)

PDF #	QM	Chemical Formula	Compound Name	Weight %	D1	D2	D3	L1 ▲	SYS	Dcalc
00-038-1974	S	C6 H16 N Cl	Triethylamine hydrochloride	C52.35 Cl25.76 H11.72 N10.18	7.220000	3.217000	3.535000	7.220000	H	1.073
02-070-0983	S	C8 H7 N O3	3-Nitroacetophenone	C58.18 H4.27 N8.48 O29.06	3.218650	7.227940	5.595800	7.227940	M	1.433
00-050-2257	O	C8 H7 N2 Na O6	Sodium 6,8-dinitro-1,4-dioxaspiro(4.5)deca-6,9-dienide	C38.41 H2.82 N11.20 Na9.19 O38.38	7.501000	3.214000	7.211000	7.501000	X	
02-070-3637	S	C10 H10 Si2	1,2-Disila-acenaphthene	C64.45 H5.41 Si30.14	7.575240	7.222260	3.212950	7.575240	M	1.3
02-071-5970	S	C12 H6 O5	2-Dehydro-2-methyl-5,8-dioxo-5,8-dihydrofuro(3,2-g)chromen-4-one	C62.62 H2.63 O34.75	3.218460	7.220580	7.452950	7.709540	M	1.607
02-062-4967	S	C10 H8 Cl2 N2 Sn	Dichloro-(2,2'-bipyridyl-N,N')-tin(ii)	C34.74 Cl20.51 H2.33 N8.10 Sn34.32	7.229820	7.860400	3.226280	8.281450	M	1.967
02-063-6684	B	(C12 H12 Mn N2 O4 P2)n	1,10-Phenanthroline-bis(phosphinato) manganese(ii)	C39.47 H3.31 Mn15.05 N7.67 O17.53 ...	7.224100	3.224160	4.359590	9.469800	M	1.743
00-020-1486	I	4(NH2 · C S · NH2) Cs I	Cesium Iodide Thiourea	C8.51 Cs23.55 H2.86 I22.49 N19.86 S...	3.230000	5.100000	7.210000	10.200000	T	

Unique result achieved with density
 OR elemental analysis
 OR a third d-spacing
 OR the long d-spacing
 OR a chemical subfile

Olazepine Form X

Is the description unique ?

YES ! – Claim 3 No reference matches all eight d-spacings to within 1%

YES ! - Claim 1 - Olazapine with a d-spacing about 11.05

YES ! – Claims 6,7,8 – with extensive d-spacing listings

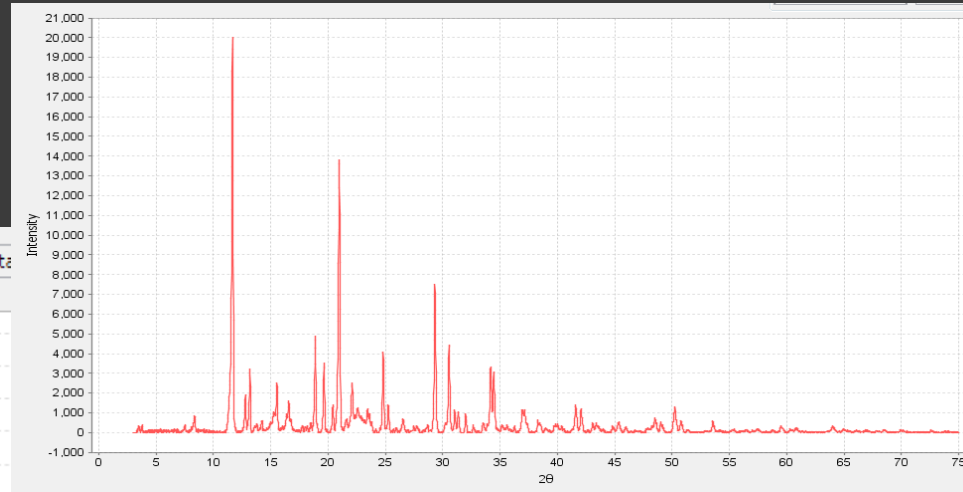
3 Critical Observations

Chemistry

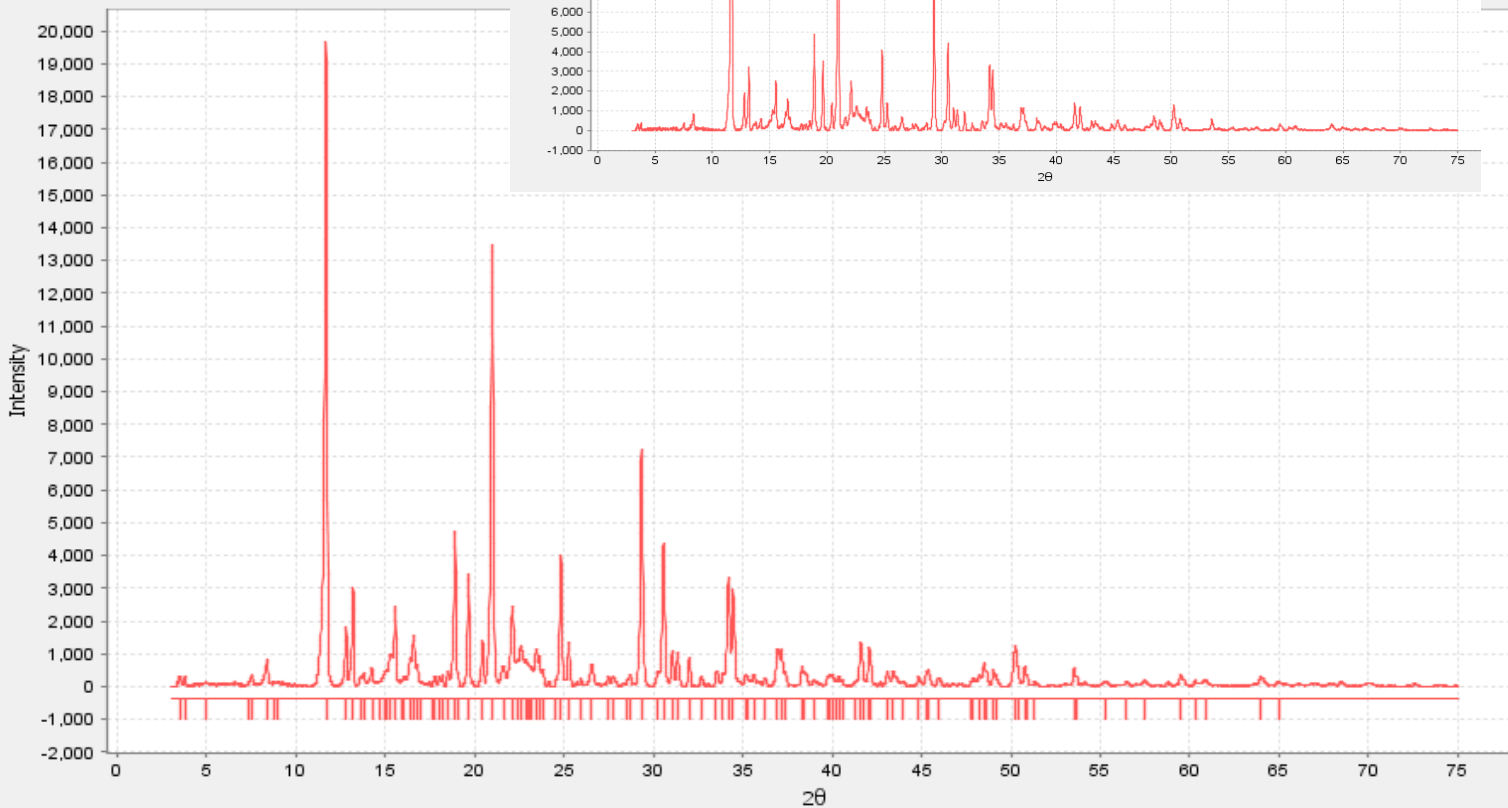
d-spacing

Estimated d-spacing error 1%

Donnatal



File Name: C:\Users\fawcett\Desktop\TEST EXAMPLES\Donnatal

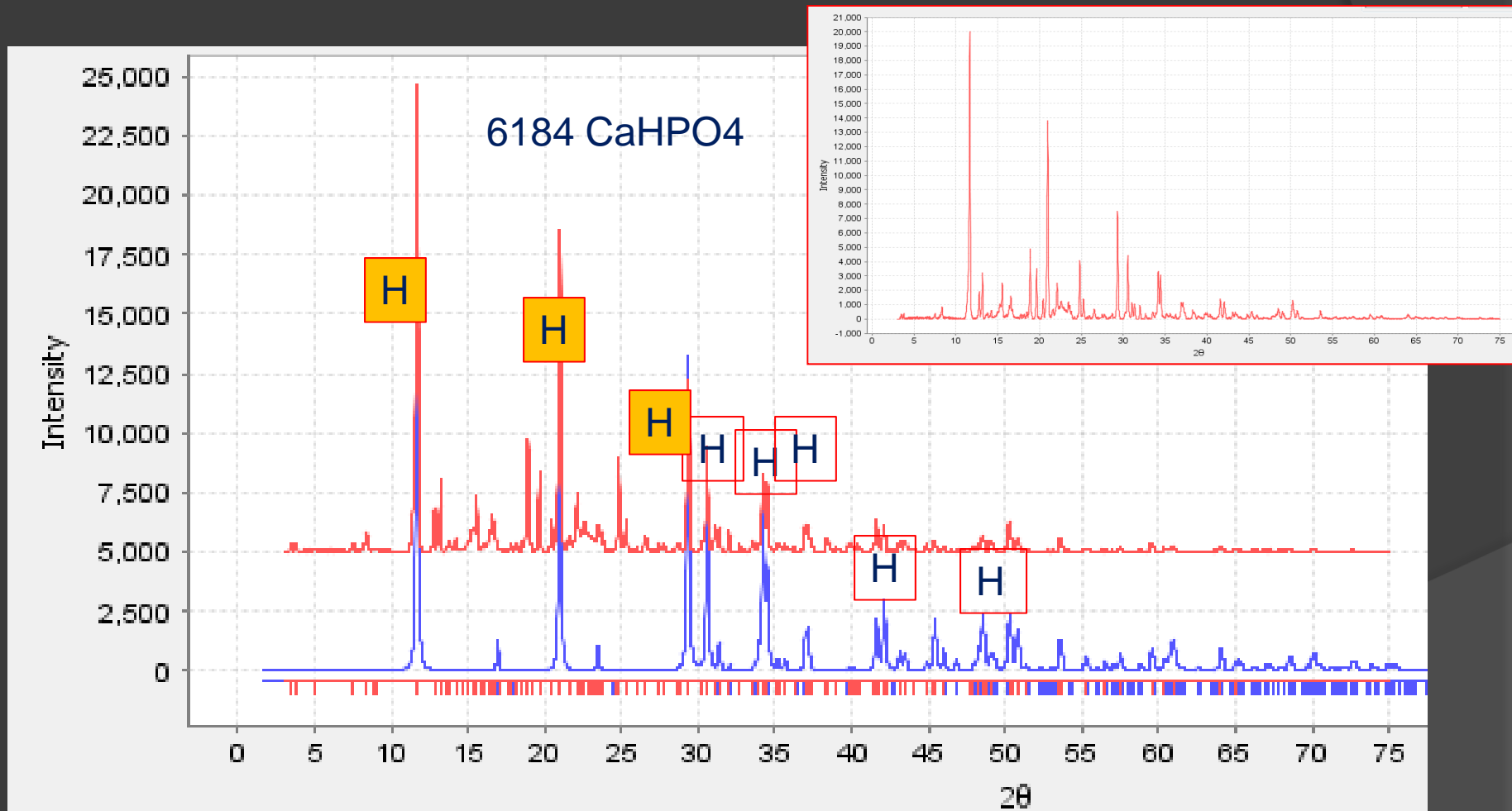


Peaks (116)

2θ	Intensity
3.5086357866	333.8477263...
3.8088132586	324.6190967...
4.976170094	117.9636461...
7.3775898696	140.5859498...
7.5110020794	342.7833700...
8.3781814428	817.2703834...
8.7617415459	145.1503771...
8.9785363868	143.2829581...
11.6968101606	19044.83639...
12.8141374174	1787.212041...
13.1976975204	3021.258510...
13.5979341497	294.2786745...
13.8147289906	417.0314897...
14.2649951985	517.1751285...
14.6485553016	163.6501657...
14.9654092997	446.1616462...
15.0321154046	483.3203620...
15.2489102455	927.1113660...
15.5490877174	2436.669811...
15.8826182418	174.1859534...
16.0327069778	295.7739961...
16.3995905547	865.8380575...
16.5997088693	1545.594537...
16.7664741315	663.3977242...
16.9999454986	177.1821686...
17.6670065474	157.3405037...
17.7670657047	313.8026696...
18.0338901242	283.6224817...

Modified Hanawalt Method

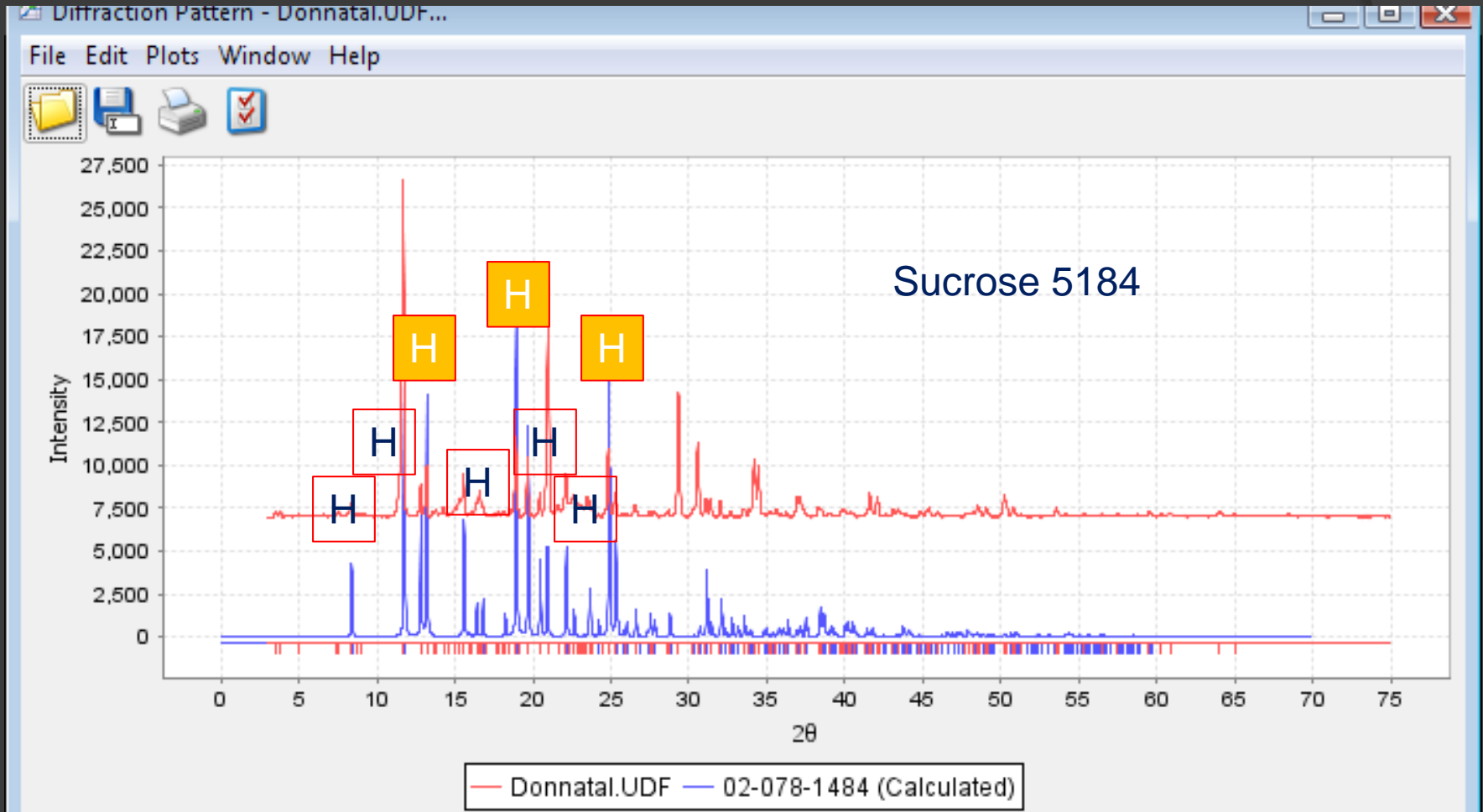
Top 8 of reference compared to the experimental
d,l list



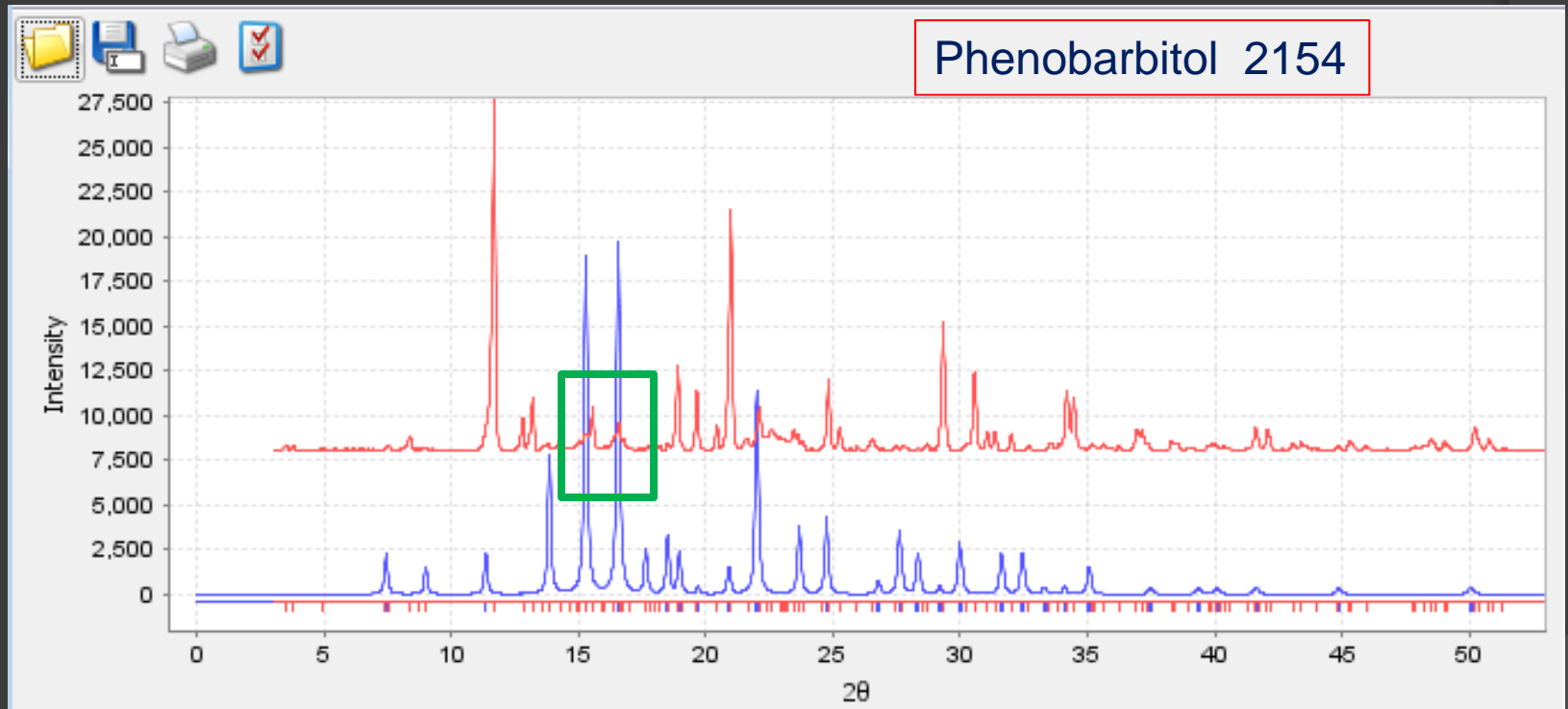
Hanawalt Method

Top 8 of reference compared to the experimental d,I list

6



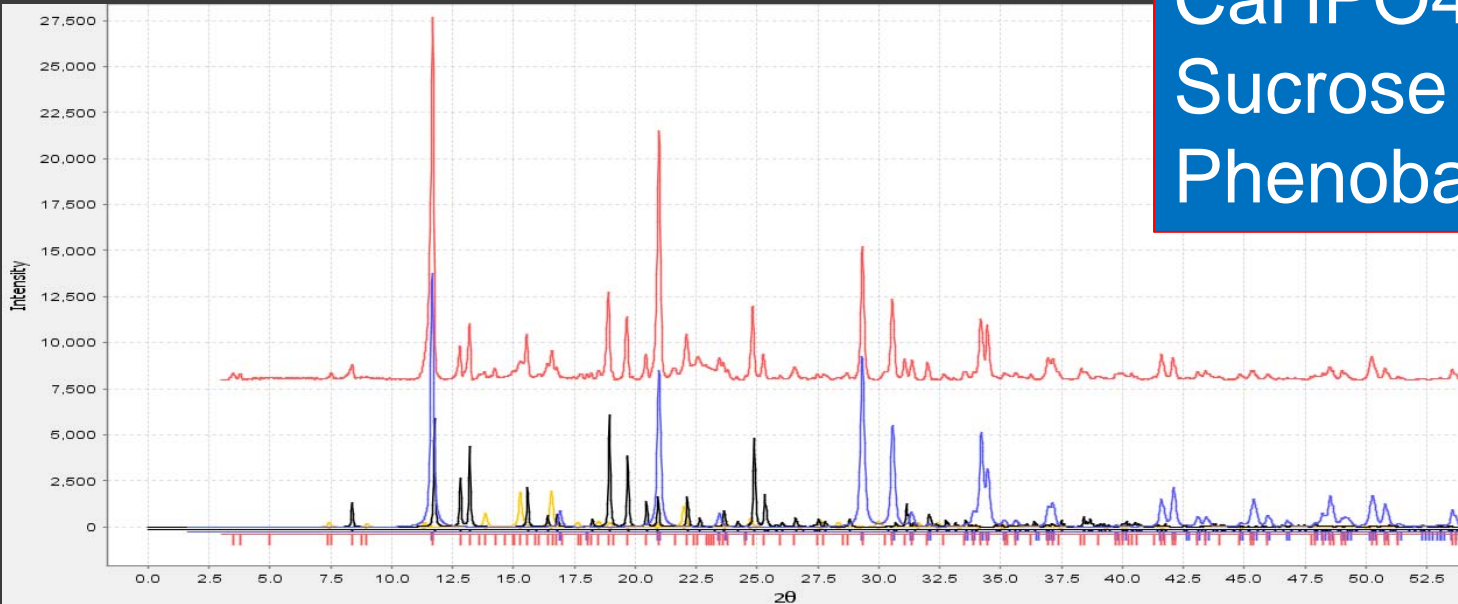
Hanawalt Method



Important Additional Information – Eight line matched, however the match was not precise resulting in a lower score. The reference was a low quality data set.

Phase Identification

CaHPO₄
 Sucrose
 Phenobarbital



Lines (90 of 116)

Ex d(Å) ▼	Ex I	P1 d(Å)	P1 I	P2 d(Å)	P2 I	P3 d(Å)	P3 I
4.90001	1						
4.91480	1	4.923200	0				
4.87014	1			4.862690	7		
4.79180	2					4.800000	16
4.69122	24			4.683920	100	4.680000	12
4.64654	1						
4.51374	17			4.506830	67	4.510000	2
4.34235	7			4.339300	24		
4.22970	70	4.237180	65	4.242170	30	4.250000	8
4.10400	3						
4.01535	12			4.012950	28	4.040000	60
3.96809	4						
3.93913	6			3.924410	9		
3.87968	4						

90/116 Peaks matched
 All intensities above 4 identified

Total Pattern Analysis

OEM Software – Total Pattern Approach

79/79 Peaks matched – Used a higher significance threshold for peak intensities

	Visible	Ref. Code	Compound Name	Chemical Formula	Score	Displ...	SUL	Semi...	TL	ML	Q
1	<input checked="" type="checkbox"/>	01-072-0713	Brushite	Ca H P O4 (H2 O...	59	-0.005	0	34	102	60	"S"
2	<input checked="" type="checkbox"/>	02-086-3216	Phenobarbital	C12 H12 N2 O3	35	-0.007	0	34	63	47	"I"
3	<input checked="" type="checkbox"/>	00-024-1977	Sucrose	C12 H22 O11	56	0.022	0	31	79	63	"S"

Matched 170 d-spacing (many overlap)
out of a total of 244 possible
Displacement 0.02 or less

All strong lines matched
(>50% I relative)

Conclusions

- 1) d-spacing accuracy is the critical parameter in phase Identification (unit cell searches depend on indexed d-spacings)
- 2) Multiple d-spacings observations greatly improves the accuracy of an identification
- 3) Examine the quality of the experimental data *and* the reference ICDD provides quality measures for all PDF references
- 4) Quality in references is continuously improving
- 5) For good quality experimental data, 3 d-spacings will usually define a unique result with the current overall quality of the Powder Diffraction File
- 6) Two d-spacings and a third independent observation (chemistry, physical property, crystal habit, subfile etc) often will lead to a unique result.

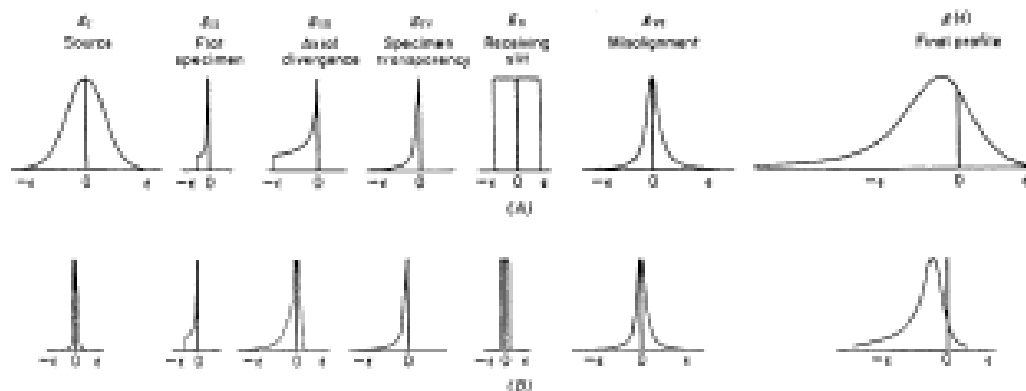


Figure 6 Graphical representation of instrumental contributions to peak profiles as depicted by Klug and Alexander.

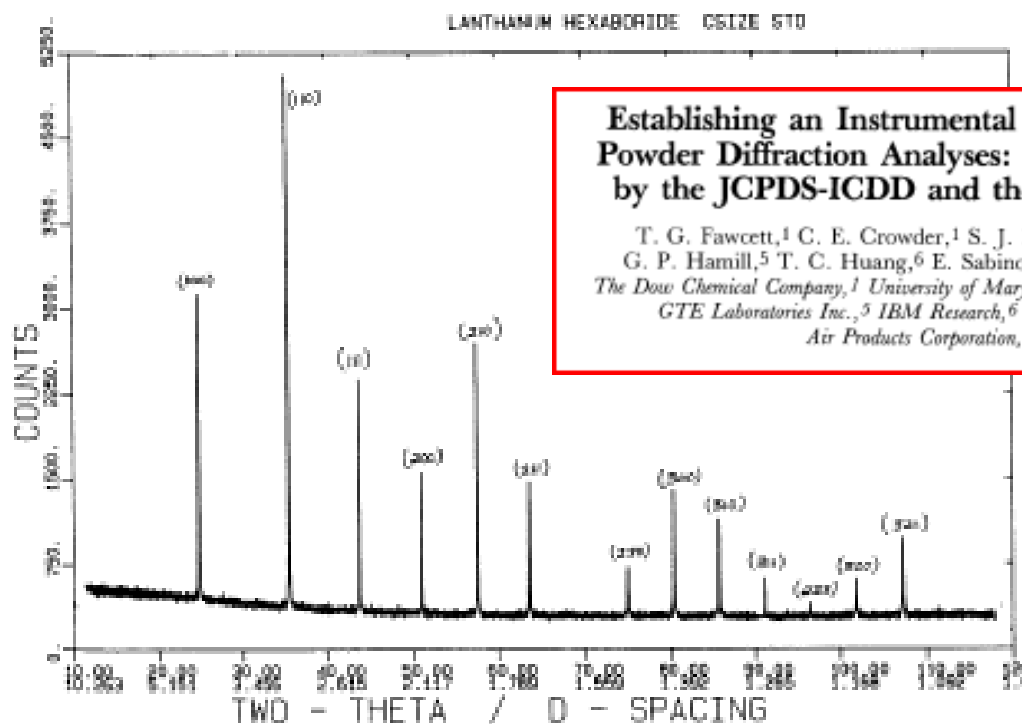


Figure 7 Representative data taken on LaB₆.

Establishing an Instrumental Peak Profile Calibration Standard for Powder Diffraction Analyses: International Round Robin Conducted by the JCPDS-ICDD and the U.S. National Bureau of Standards†

T. G. Fawcett,¹ C. E. Crowder,¹ S. J. Brownell,¹ Y. Zhang,² C. Hubbard,^{3*} W. Schreiner,⁴ G. P. Hamill,⁵ T. C. Huang,⁶ E. Sabino,⁷ and J. I. Langford,⁸ R. Hamilton,⁹ and D. Louër¹⁰
 The Dow Chemical Company,¹ University of Maryland,² The National Bureau of Standards,³ Philips Laboratories,⁴ GTE Laboratories Inc.,⁵ IBM Research,⁶ The PQ Corporation,⁷ University of Birmingham,⁸ U.K., Air Products Corporation,⁹ and the University of Rennes,¹⁰ France

NIST SRM 660

Form X – Geneva Pharmaceuticals Inc

Title: **Olanzapine crystal modification**

Document Type and Number: United States Patent 6740753

Abstract: A novel crystal form the pharmaceutical compound olanzapine, processes for its preparation and its pharmaceutical use are disclosed.

Claims: We claim:

1. Form X olanzapine characterized by a melting point in the range from 187° C. to 191° C.
2. Form X olanzapine characterized by a small peak at about 11.05 d-spacing units in its powder x-ray diffraction pattern.
3. Form X olanzapine of claim 1 having a powder x-ray diffraction pattern with characteristic peaks at about 11.05, 9.98, 6.24, 6.13, 3.75, 3.61, 3.53, 3.43 and 2.67 d-spacing units.
4. Form X olanzapine of claim 3 having a melting point of 187° C. to 190° C.
5. Form X olanzapine of claim 4 having a melting point range of 189° C. to 190° C.
6. Form X olanzapine of claim 5 having a powder x-ray diffraction pattern with characteristic peaks at about 11.05, 9.98, 6.24, 6.13, 3.75, 3.61, 3.53, 3.43 and 2.67 d-spacing units.
7. Form X olanzapine of claim 4 having a powder x-ray diffraction pattern with characteristic peaks at about 11.05, 9.98, 6.24, 6.13, 4.83, 4.71, 4.57, 4.48, 4.39, 4.32, 3.84, 3.75, 3.61, 3.53, 3.43, 2.95, 2.86, 2.67, 2.43 and 2.36 d-spacing units.
8. Form X olanzapine of claim 5 having a powder x-ray diffraction pattern with characteristic peaks at about 11.05, 9.98, 6.24, 6.13, 4.83, 4.71, 4.57, 4.48, 4.39, 4.32, 3.84, 3.75, 3.61, 3.53, 3.43, 2.95, 2.86, 2.67, 2.43 and 2.36 d-spacing units.

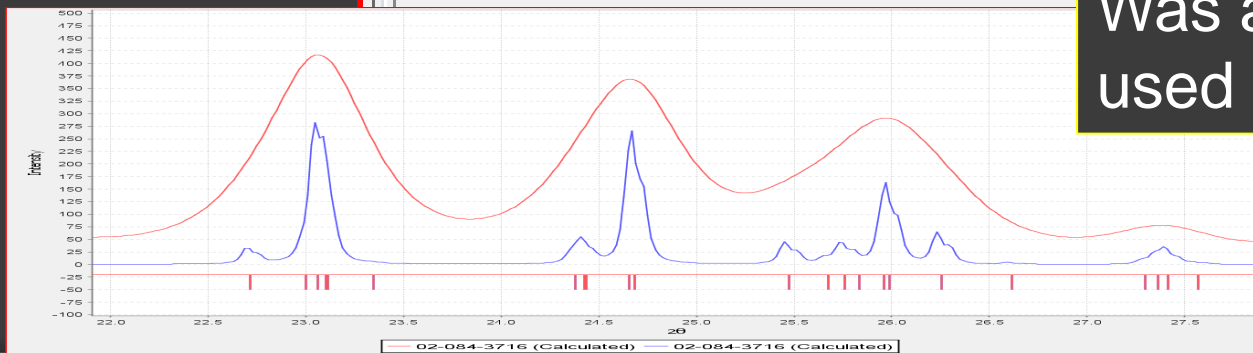
Claim 3 - What does about 11.05 mean ? (1% or 0.1 d-spacing)

Lines (6 of 9)

Ex d(Å) ▼	Ex I	P1 d(Å)	P1 I
11.05	10		
9.98	100	10.048100	100
6.24	10	6.225530	9
6.13	10	6.141700	6
3.75	10		
3.61	10	3.606510	32
3.53	10		
3.43	10	3.428600	20
2.67	10	2.666020	3

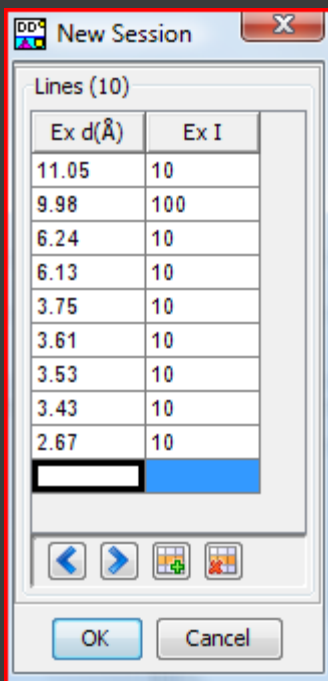
Olanzapine methanol monohydrate

Was an Internal Std.
used to calibrate 11.05 !



Is the description unique ?

Claim 3



The screenshot shows a 'New Session' dialog box with a table titled 'Lines (10)'. The table has two columns: 'Ex d(Å)' and 'Ex I'. The data is as follows:

Ex d(Å)	Ex I
11.05	10
9.98	100
6.24	10
6.13	10
3.75	10
3.61	10
3.53	10
3.43	10
2.67	10

No chemistry

Applied esd
Two theta

0.18
0.09
0.01

Match Results on 8 d-spacings

355 Compounds matched 2-8 d-spacings
51 Compounds matched 2-8 d-spacings
14 Compounds matched 2-8 d-spacings

Olanzapine Subfile

Applied esd
Two theta

0.18
0.09
0.01

Match Results on 8 d-spacings

5 Compounds match 2-8 d-spacings
5 Compounds match 2-8 d-spacings
1 Compound matches a single peak

Patent Claims with 1-21 peaks

*Observed Peaks – 1% error, no intensity information
No Chemistry*

1) 11.05 (0.11) **11,150 Entries**

2) 9.98 (0.1) 15,727 Entries

3) 6.24 (0.06) 15,630 Entries

Present in all claims
using 1 peak and 1% accuracy

1 & 2 & 3

2 Entries Match using all materials and ~ 1% error
(goes to 20 with slightly larger error)

4) 6.14 (0.06)

1 & 2 & 3 & 4

2 Entries Match
(53 with slightly larger error)

Extra data
didn't help

Polymorph X is not in the
database

Chemistry is essential to a
unique claim